Measurement of the $B^0\overline{B}^0$ Oscillation Frequency Using Fully-Reconstructed Semileptonic Decays

by

Christopher Mark LeClerc

A.B. (University of California-Berkeley) 1991

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Physics

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Robert Jacobsen, Chair
Professor Steven Derenzo
Professor Lawrence Hall

Fall 2002

The dissertation of Christopher Mark LeClerc is approved:

_____

Chair   Date

_____

                                                                Date

_____

                                                                Date

University of California at Berkeley

Fall 2002

Abstract

Measurement of the $B^0\overline{B}^0$ Oscillation Frequency Using

Fully-Reconstructed Semileptonic Decays

by

Christopher Mark LeClerc

Doctor of Philosophy in Physics

University of California at Berkeley

Professor Robert Jacobsen, Chair

$B^0\overline{B}^0$ flavor oscillation frequency is studied by examining events containing a fully reconstructed $B \rightarrow D^* l\nu$. The opposite side flavor is determined with tagging tools that look for characteristic kaon and lepton tags, plus a neural network output for events with conflicting or non-existent kaon/lepton tags. Based on a data set of 20.6 fb^{-1} on the $\Upsilon(4s)$ resonance, 2.6 fb^{-1} offpeak, Δm_d = 0.492 $\pm$ 0.018 $\pm$ 0.013 ps^{-1} . The first error is statistical and the second is systematic. The B^0 lifetime is also measured, τ_{B^0} = 1.523 $^{+0.024}_{-0.023} \pm 0.022$ ps. The Δm_d measurement is competitive with other recent measurements and develops several novel and powerful data analysis techniques.

I dedicate my thesis to my loving wife, whose love and support means everything to me.

Contents

Acknowledgements

I want to thank the follow people for the extensive help and advice provided:

My advisor: Bob Jacobsen

Co-workers: Chih-hsiang Cheng, Tim Meyer and David Kirkby

Faculty and staff: Natalie Roe, Robert Cahn, Patricia Burchat

Special thanks to Mandeep Gill who is both a co-worker and provided exceptional support helping me finish on time!

Chapter 1

Introduction

"There is nothing new to be discovered in physics now. All that remains is more and more precise measurement." - Lord Kelvin [1][1]

Today's physicist most likely has a very different world view than Lord Kelvin. Unlike Kelvin who asserted that there were no major advances left to make in physics, many physicists today actively seek to unseat the currently accepted and spectacularly successful "Standard Model of Particle Physics". There is great anticipation for the next generation of accelerators (the accelerators at Fermilab and CERN) to possibly discover "physics beyond the Standard Model". Even current experiments like Belle and BaBar have the potential of uncovering something new.

1.1 Standard Model

The Standard Model has great computational powers. It can be used to calculate particle lifetimes, reaction rates, cross sections, etc... and can do so with tremendous accuracy. The weaknesses of the Standard Model include the fact that the Standard Model

---

[1]1900 at the British Association for the Advancement of Science.

neglects gravity and also that it critically depends on a set of 18 parameters [2], which all must be measured, not derived from basic principles. These 18 parameters include the 6 quark masses, 3 lepton masses, 4 elements from quark mixing (CKM elements), 2 parameters involving the yet to be discovered Higgs (the Higgs mass and vacuum expectation value) and 3 parameters which describe the coupling strengths of the strong, weak and electromagnetic forces. The discovery of neutrino mass will inevitably expand this number of parameters by at least seven (3 neutrino masses and 4 mass mixing parameters - 6 mixing parameters if the neutrino is a majorana particle). This leads many physicists to believe that the Standard Model is possibly an approximation of a larger, more encompassing theory, one that includes gravity and has only a few parameters. Much of the theoretical framework of Standard Model was in place during the 1970's. The confirmation of the Standard Model with remarkable precision occurred during the 1980's and 1990's [4]. The rest of this chapter is devoted to describing the Standard Model. The rest of the paper is devoted to an experiment which hopes to discover deviations from the "Standard Model".

1.1.1 Elementary Particles - quarks, leptons and forces

The Standard Model attempts to describe all the particles and interactions which make up our universe. The particles and their interactions are concisely described by listing the particle's quantum numbers - fundamental attributes intrinsic to elementary particles. The relevant quantum numbers include the particle's mass, charge, "color charge", spin and flavor. All fundamental particle masses (except the ones identically equal to zero) are given by measuring 11 of the 18 free parameters of the Standard Model. The charge describes a particles coupling with the electromagnetic force while flavor is im-

portant for the weak force. Color charge determines the strong interactions and the spin describes the type of statistics the particle will obey [5].

From quantum mechanics, all particles can be broken into two classes based on their spin. Particles with half-integer units of spin are called Fermions and they obey Fermi statistics. The most important and familiar result from Fermi statistics is the Pauli Exclusion Principle. This principle states that no two identical fermions can occupy the same state at the same time, i.e. have the exact same quantum numbers. Particle with integer spins are called bosons. Not only can two identical bosons occupy the same state at the same time, multiply occupied states are favored over singly occupied states.

Table 1.1 shows leptons and quarks that make up matter. All the quarks and leptons are fermions. All the force carrying particles in Table 1.2 are bosons.

Leptons (spin=1/2 $\hbar$)			Quarks (spin=1/2 $\hbar$)		
Flavor	Mass (GeV/c^2)	Charge	Flavor	Mass (GeV/c^2)	Charge
ν_e	$<3\times10^{-9}$	0	u up	0.001-0.005	2/3
e	0.000511	-1	d down	0.003-0.009	-1/3
ν_μ	$<2\times10^{-4}$	0	c charm	1.15-1.35	2/3
μ	0.106	-1	s strange	0.075-0.170	-1/3
ν_τ	$<2\times10^{-2}$	0	t top	174.3$\pm$5.1	2/3
τ	1.777	-1	b bottom	4.0-4.4	-1/3

Table 1.1: Characteristics of Fermions in the Standard Model broken down by 1st, 2nd and 3rd generations [6].

There are three generations or families of particles. Each generation has two quarks, one with charge +2/3 (an "up" type quark) and one with -1/3 ("down" type quark), a negatively charged lepton and a chargeless neutrino. Each member of the family has an antimatter partner that has the same spin and mass, but opposite charge and flavor quantum numbers. The family members get progressively heavier as one proceeds from the first to second to third generation. These particles are currently considered fun-

Interaction	Gravity	Weak	Electromagnetic	Strong
Acts on Particles	Mass-Energy All	Flavor Quarks, Leptons	Electric Charge Electrically Charged	Color Charge Quarks, Gluons
Force carrying Boson	Graviton	W^+, W^-, Z^0	γ	Gluons $(g_1,...,g_8)$
Spin	$2\hbar$	$\hbar$	$\hbar$	$\hbar$
Charge	0	$\pm 1,0$	0	0
Mass (GeV)	0	$\pm 80.4, 91.2$	0	0

Table 1.2: Characteristics of Force carrying Bosons in the Standard Model [6].

demental, although its possible they are made up of smaller entities that only reveal themselves at extraordinarily high energies.

The four forces of nature all have associated particles which mediate all interactions. The graviton is a spin-2 particle which has yet to be detected. The only other undetected particle accepted in the Standard Model is the Higgs Boson. The graviton is assumed to be massless because of its long range. Only massless particles produce long range inverse square force laws. The long range nature of electromagnetism implies that the photon is also massless. There are several striking differences between electromagnetism and gravity. The first is that gravity is purely attractive while electromagnetism is attractive and repulsive. The second is their relative strengths. If we compare their dimensionless coupling constants, $G_N M^2/\hbar c \approx 6.7 \times 10^{-39} (GeV^2) \times M^2$ for gravity and $\alpha = e^2/\hbar c \approx 1/137$ for electromagnetism [6], we find that gravity is roughly equal to electromagnetism for a particle with unit charge and a mass $Mc^2 \approx 10^{18}$ GeV. Today's largest particle accelerators reach energies $\sim 10^3$ GeV, the Large Hadron Collider (LHC) will reach $\sim 10^4$ GeV around 2007. The gravitational attraction between two protons is $\sim 10^{36}$ times smaller than the electromagnetic repulsion. The disparity is much larger for lighter particles like the electron or light quarks. This vast difference in relative strength makes

gravity all but irrelevant on the subatomic level until you reach energies approaching the Planck scale (the energy where all forces are expected to be equally important $\approx 10^{19}$ GeV).

The weak force is mediated by the W and Z bosons. The large masses associated with these particles means that the force can only act over a short range, roughly estimated by the inverse of the force carrying particle's mass. The weak force and electromagnetism can be combined into a single electroweak force which contains four gauge bosons and is described by the SU(2)$\times$U(1) symmetry group. The coupling strength for both the weak force and the electromagnetic force are significantly smaller than unity, which allows for perturbative calculations [5]. The electroweak force will be discussed further in section 1.3.

The strong force is described by Quantum Chromodynamics with the SU(3) symmetry group. Each quark has a "color charge" with one of three values, red(r), green(g), or blue(g). An anti-quark carries anti-color, $\overline{r}$, $\overline{g}$, or $\overline{b}$. Gluons carry and transfer color charge. A standard representation for the eight different gluons is, $r\overline{b}$, $r\overline{g}$, $b\overline{r}$, $b\overline{g}$, $g\overline{r}$, $g\overline{b}$, $(r\overline{r}\text{-}b\overline{b})$, $(r\overline{r}+b\overline{b}\text{-}2g\overline{g})$. Only quarks and gluons have color, so only these particles participate in strong interactions. At high energies (at the Z mass of 91 GeV), the strong coupling constant is relatively small (0.1) compared to unity and perturbative calculations can be done. However, at lower energies the coupling constant gets stronger making secondary interactions more important thus invalidating any perturbative calculations. This also leads to color confinement, which states that particles can only be colorless (color singlet) so no free quarks or gluon can be found. When quarks in a baryon or meson are separated, during a collision in an accelerator for example, it is energetically favorable for a correctly colored quark-antiquark pair to appear from the vacuum and "cover" the bare

quarks. So far, quarks have only been found as baryons or mesons [5].

1.1.2   Composite Particles - Baryons and Mesons

A baryon is a three quark color singlet state (rgb) and a meson is a quark-antiquark state ($color - \overline{color}$). There are six different quark flavors listed in table 1.1. If you consider n different flavors, there exists an approximate $SU(n)$ flavor symmetry. This symmetry is broken by the different quark masses. The first two quarks have very similar masses, and the SU(2)$_{flavor}$ symmetry (isospin) is a very good symmetry - as shown by the small mass difference between protons and neutrons. The strange quark is modestly heavier than the up and down quarks, giving an approximate SU(3)$_{flavor}$ symmetry for (u,d,s) quarks and all higher SU(n)$_{flavor}$ symmetries are badly broken by the heavy quarks.

Baryons are fermions, which by definition have an antisymmetric wave function. The SU(3) color singlet state is a completely antisymmetric state, so a baryon wave function can be decomposed as follows:

$$|qqq> = |color singlet>_A \times |spin, flavor, space>_S \qquad (1.1)$$

The color quantum number has a large impact on the types of baryons that can be formed. Without the antisymmetric color component, the Δ^{++} and other baryons with three identical quarks would be impossible to form since the space, spin and flavor part of the wave function would be symmetric [5].

The baryons that can be constructed from the lightest 3 quarks obey the approximate SU(3)$_{flavor}$ symmetry and belong to multiplets described by:

$$3 \otimes 3 \otimes 3 = 10_S \oplus 8_M \oplus 8_M \oplus 1_A \qquad (1.2)$$

SU(3)	quark content	SU(3)	quark content
n, p	udd,uud	$\Delta^-, \Delta^0, \Delta^+, \Delta^{++}$	ddd,udd,uud,uuu
Σ^-, (Σ^0, Λ^0), Σ^+	(dd,ud,uu)$\times$s	$\Sigma^-, \Sigma^0, \Sigma^+$	(dd,ud,uu)$\times$s
Ξ^-, Ξ^0	(d,u)$\times$ss	Ξ^-, Ξ^0	(d,u)$\times$ss
		Ω^-	sss
Σ_c^0, $(\Sigma_c^+, \Lambda_c^+)$, Σ_c^{++}	(dd,ud,uu)$\times$c	$\Sigma_c^0, \Sigma_c^+, \Sigma_c^{++}$	(dd,ud,uu)$\times$c
$\Xi_c^0(2)$, $\Xi_c^+(2)$	(d,u)$\times$sc	Ξ_c^0, Ξ_c^+	(d,u)$\times$sc
Ω_c^0	ssc	Ω_c^0	ssc
$\Xi_{cc}^+, \Xi_{cc}^{++}$	(d,u)$\times$cc	$\Xi_{cc}^+, \Xi_{cc}^{++}$	(d,u)$\times$cc
Ω_{cc}^+	scc	Ω_{cc}^+	scc
		Ω_{ccc}^{++}	ccc

Table 1.3: Partial list of baryons in ground state SU(4) multiplet [6]. The first row contains uncharmed baryons, the second row lists baryons with one charm quark, the third row lists baryons with two charm quarks while the last row contains baryons with 3 qurks.

Where the S, M and A mean symmetric, mixed or antisymmetric under the interchange of any two quarks. Table 1.3 shows baryons extended to SU(4), anything made with u,d,s,c quarks. The first column is a 20-plet based on a SU(3) octet. The second column of baryons is based on the totally symmetric SU(3) decuplet. A similar extension can be done to include bottom quarks in an approximate SU(5).

The mesons are inherently simpler since there are only two quarks and they are not identical - they are a quark/antiquark pair. With just two quarks, the meson can have a spin S of 0 or 1. If the orbital angular momentum, L, is zero, all mesons will be either pseudoscalar particles, J^P=0^-, or vector particles, J^P=1^-. For the 3 lightest quarks, the multiplets are described by:

$$3 \otimes \overline{3} = 8 \oplus 1 \qquad (1.3)$$

Both the pseudoscalar and vector mesons will form an octet plus singlet. Again this can be extended for the charm and bottom quarks, as shown in table 1.4

States with the same quantum numbers can mix, so the $I = 0$ state of the octet

	quark content	pseudoscalar mesons	vector mesons
u,d,s Mesons	$u\overline{d}$, $(u\overline{u}\text{-}d\overline{d})$, $d\overline{u}$	π^+, π^0, π^-	ρ^+, ρ^0, ρ^-
	$(u\overline{u}+d\overline{d})$, $s\overline{s}$	η', η	ω, ϕ
	$\overline{s}u$, $s\overline{u}$, $\overline{s}d$, $s\overline{d}$	$K^+, K^-, K^0, \overline{K}^0$	$K^{*+}, K^{*-}, K^{*0}, \overline{K}^{*0}$
Charmed Mesons	$c\overline{d}$, $\overline{c}d$, $c\overline{u}$, $\overline{c}u$	$D^+, D^-, D^0, \overline{D}^0$	$D^{*+}, D^{*-}, D^{*0}, \overline{D}^{*0}$
	$c\overline{s}$, $\overline{c}s$	D_s^+, D_s^-	D_s^{*+}, D_s^{*-}
	$c\overline{c}$	η_c	J/ψ
Bottom Mesons	$\overline{b}u$, $b\overline{u}$, $\overline{b}d$, $b\overline{d}$	$B^+, B^-, B^0, \overline{B}^0$	$B^{*+}, B^{*-}, B^{*0}, \overline{B}^{*0}$
	$\overline{b}s$, $b\overline{s}$	$B_s^0, \overline{B}_s^0$	$B_s^{*0}, \overline{B}_s^{*0}$
	$b\overline{b}$	η_b	$\Upsilon(1S)$

Table 1.4: Pseudoscalar and vector mesons [6].

can mix with the singlet state. For the pseudoscalars, this produces the η and η' [5].

$$\eta = \eta_8 \cos\theta - \eta_1 \sin\theta \qquad (1.4)$$

$$\eta' = \eta_8 \sin\theta + \eta_1 \cos\theta \qquad (1.5)$$

where the octet and singlet are defined as:

$$\eta_1 = \frac{1}{\sqrt{3}}(u\overline{u} + d\overline{d} + s\overline{s}) \qquad (1.6)$$

$$\eta_8 = \frac{1}{\sqrt{6}}(u\overline{u} + d\overline{d} - 2s\overline{s}) \qquad (1.7)$$

The final mixed state is approximately:

$$\eta = \frac{1}{\sqrt{3}}(u\overline{u} + d\overline{d} - s\overline{s}) \qquad (1.8)$$

$$\eta' = \frac{1}{\sqrt{6}}(u\overline{u} + d\overline{d} + 2s\overline{s}) \qquad (1.9)$$

The same is true for vector mesons also. The final mixed states for the ϕ and ω are approximately:

$$\phi = s\overline{s} \qquad (1.10)$$

$$\omega = \frac{1}{\sqrt{2}}(u\overline{u} + d\overline{d}) \qquad (1.11)$$

1.1.3 Particle decays

Most baryons and mesons decay into lighter elementary particles with relatively short lifetimes. Particles can decay via the strong, weak or electromagnetic forces. The characteristic lifetimes are generally determined by the force mediating the decay, but the kinematics of the decay and various symmetries can also greatly affect the result. The typical lifetime of a strong decay is 10^{-23} sec, electromagnetic is 10^{-16} sec and weak decay is 10^{-8} sec. The only stable particles are the proton, electron, photon and neutrinos. All baryons will ultimately decay into a proton plus extras (photons, leptons and neutrinos). All mesons eventually decay into leptons and neutrinos. The τ and μ leptons follow a similar decay chain, always leaving an electron plus neutrinos and possibly photons.

The decay of any particle can be neatly broken into the product two pieces, one represents the pure kinematics of the events and the other is the matrix element between the initial and final states which carries the information about the physics involved in the particular decay. For the decay $X \rightarrow 1 + 2 + ... + n$: [4]

$$d\Gamma = \frac{1}{2E_X}|\mathcal{M}|^2 \frac{d^3 p_1}{(2\pi)^3 2E_1} ... \frac{d^3 p_n}{(2\pi)^3 2E_n} (2\pi)^4 \delta^{(4)}(p_X - p_1 - ... - p_n) \qquad (1.12)$$

Where $|\mathcal{M}| = <f|\mathcal{H}|i>$ is the matrix element. For the special case of a two body decay,

$$\Gamma(X \rightarrow 1 + 2) = \frac{p_f}{32\pi^2 m_X^2} \int |\mathcal{M}|^2 d\Omega \qquad (1.13)$$

All angular dependence is carried by the matrix element $\mathcal{M}$. The final momentum of both daughters is equal in the rest frame of the particle X and is denoted by p_f. The total decay width of a particle is the sum of all individual channels, [4]

$$\Gamma_{total} = \sum \Gamma_i \qquad (1.14)$$

The decay rate is related to a particles lifetime through the equation

$$N_X(t) = N_X(0)e^{-\Gamma t}, \text{ thus } \tau = 1/\Gamma \tag{1.15}$$

Table 1.5 shows the mass, lifetime and decays of some particles which are particularly relevant to this analysis. One interesting thing you can gather from the Table is which particles we can expect to see travel from a parent's decay vertex and which ones decay almost instantaneously. The SVT resolution is on the order of tens of microns (μm). We should clearly be able to see the B^0 and D^0 fly away from the parent and produce a separate decay vertex. The D^* decays so fast that the daughters of the D^* parent, the D^*'s pion and D meson all appear to come from the same vertex.

meson	mass (MeV)	lifetime (sec)	$c\tau$ (μm)	Daughters	Γ_i/Γ %
$\pi^\pm$	139.57	2.6×10^{-8}	7.8×10^6	$\mu^\pm\nu_\mu$	99.99
π^0	134.98	8.4×10^{-17}	0.025	2γ	98.80
$K^\pm$	493.68	1.2×10^{-8}	3.7×10^6	$\mu^\pm\nu_\mu$	63.51
				$\pi^\pm\pi^0$	21.16
K_S	497.67	0.89×10^{-10}	2679	$\pi^+\pi^-$	68.61
				$\pi^0\pi^0$	31.39
D^0	1864.5	4.1×10^{-13}	123.7	$K^-\pi^+$	3.83
				$K^-\pi^+\pi^-\pi^+$	7.49
				$K^-\pi^+\pi^0$	13.9
				$K_S\pi^+\pi^-$	5.4
D^{*+}	2010.0	$<5\times10^{-21}$	$<1.5\times10^{-6}$	$D^0\pi^+$	67.7
				$D^+\pi^0$	30.7
B^0	5279.4	1.548×10^{-12}	464	$D^{*-}l^+\nu$	4.6

Table 1.5: Particle masses and lifetimes useful for this analysis [6]. The daughters column shows either the primary decay or a decay which is used in this thesis.

1.2 CP violation

There are three discrete operators that are potential symmetries of the Standard Model Lagrangian. Two operators are space-time symmetries, Parity(P) and Time rever-

sal(T). The Parity operation takes all space coordinates and replaces them with oppositely signed quantities - (x,t)→(-x,t). This has the effect of reversing the handedness of a particle, similar to viewing the particle's mirror image. The time reversal operator replaces the time coordinate with an oppositely sign value - (x,t)→(x,-t), reversing the direction of time. This is similar to watching a film of a particular event backwards. The third operation is Charge Conjugation(C), which simply replaces all particles with their anti-particles (and vice-versa). [7]

The product of all three operations (CPT) is a perfect symmetry of any local field theory by construction: you simply can't create a renormalizable local field theory without CPT being conserved. The individual symmetries are conserved by both the strong and electromagnetic forces. Initially, physicists had no reason to suspect that these symmetries were violated by the weak force, but in 1956, Lee and Yang proposed that the weak force violated parity to explain the $K \to 2\pi$ and 3π, two states with opposite parity. Immediately following this, several new experiments were conducted to confirm this, the first of which was a study of the decay of ^{60}Co with the nuclear spins lined up parallel in a magnetic field: [5]

$$^{60}Co \to^{60} Ni^* + e^- + \overline{\nu}_e \tag{1.16}$$

This decay changes the total angular momentum of the nucleus by 1 unit of angular momentum, which must be carried away by the electron and anti-neutrino. The anti-neutrino is always right handed which means the electron must be left handed to conserve angular momentum. Since the charged weak current only couples to the right handed anti-neutrino and left handed electron, there is no "mirror image" reaction. There have been no left handed anti-neutrinos observed in nature, so the parity operator symmetry is maximally broken in this case. The fact that there are left handed neutrinos

(right handed anti-neutrinos) but no left handed anti-neutrinos (right handed neutrinos) violates the charge conjugation operator. The combination of parity and charge conjugation (CP) seemed to be a good symmetry since CP takes a left handed neutrino into a right handed anti-neutrino, but even this symmetry was shown to be broken in the kaon system in 1964 by James Cronin and Val Fitch at Brookhaven National Laboratory. CP violation is a very small effect, unlike the maximal symmetry breaking of parity and charge conjugation by the weak force. [5]

CP violation has several important implications. The most dramatic is that CP violation represents a difference in the decay rates of matter vs anti-matter which offers a possible explanation for the matter/anti-matter asymmetry currently observed in the universe. Without an asymmetry in the amount of matter vs antimatter in the universe, all the matter and anti-matter should have ultimately annihilated each other while the universe cooled as it expanded. The resulting universe would be void of all matter and be filled solely with photons. [8]

Another implication of CP violation is T violation. Since the combination of CPT must be a perfect symmetry, T must be violated in such a way that that CP violation doesn't spoil the conservation of CPT. This implies that in certain cases you can find subatomic processes which act differently in the forward time direction than in the backward time direction.

1.3   Electroweak Interactions

Many of the decays in this analysis and the mixing phenomenon are governed by the weak force. The weak force and electromagnetism are united in a single theory in which the electroweak Lagrangian is based on the $SU(2) \times U(1)$ symmetry group contain-

ing four bosons. Above the spontaneous symmetry-breaking threshold, all four bosons $(W^\pm, W^0, B^0)$ are massless and the SU(2)×U(1) symmetry is unbroken. [5]

$$\mathcal{L}_{electroweak} = g(J^i)^\mu W_\mu^i + \frac{g'}{2}(j^Y)^\mu B_\mu \tag{1.17}$$

J^i is the weak isospin current and j^Y is the weak hypercharge current.

Below the symmetry-breaking threshold, the charged bosons and one linear combination of the neutral bosons acquire mass $(W^\pm, Z^0)$ while the orthogonal combination of neutral bosons remains massless (the photon). The SU(2)×U(1) symmetry is broken, but not completely. There remains a U(1) symmetry which gives rise to electromagnetism and the massless photon. The $W^\pm, Z^0$ and photon can be described in terms of the previous fields by:

$$W_\mu^\pm = \sqrt{\frac{1}{2}}(W_\mu^1 \mp iW_\mu^2) \tag{1.18}$$

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \tag{1.19}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W \tag{1.20}$$

where θ_W is the weak mixing angle, the angle which produces the mass eigenstates of the photon and Z^0.

The complete Lagrangian can now be written with the more familiar $W^\pm$, Z^0 and the photon (new weak isospin notation, $J_\mu = \frac{1}{2}(J_\mu^1 + iJ_\mu^2)$).

$$\mathcal{L}_{electroweak} = \frac{g}{\sqrt{2}}(J^\mu W_\mu^+ + J^{\mu\dagger}W_\mu^-) + ej_\mu^{em} A^\mu + \frac{g}{\cos\theta_W}J_\mu^{NC} Z^\mu \tag{1.21}$$

This equation makes use of the following relations:

$$e = g\sin\theta_W = g'\cos\theta_W \tag{1.22}$$

$$ej_\mu^{em} = e(J_\mu^3 + \frac{1}{2}j_\mu^Y) \tag{1.23}$$

$$J_\mu^{NC} = J_\mu^3 - \sin^2\theta_W j_\mu^{em} \tag{1.24}$$

Another aspect of the SU(2)×U(1) symmetry group is the distinction the weak force makes between left and right handed particles. The charged current (CC) interactions proceed only between left handed particles, while the neutral current (NC) interactions are predominantly left handed (right handed particles participate, but only through the j_μ^{em} portion of the NC). The left handed particles are contained in particle doublets, while right handed particles form singlet states.

The weak eigenstates of the quarks differs from the mass eigenstates. It is necessary to describe the down type weak eigenstate quarks as a superposition of down type mass eigenstate quarks. One could have used the up type quarks instead of down type, but there would be no difference since the absolute phases of quark wave functions isn't observable. More complicated mixing schemes can be employed but can always be simplified down to this choice of phases. The quark mixing is given by the Cabbibo-Kobayashi-Maskawa (CKM) mixing matrix: [5]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.25}$$

so the charged current only acts on the left handed doublets shown in equation 1.26, (u=up type, d=down type, l=lepton). Both left and right handed particles can interact with the photon and Z^0.

$$\phi_L = \begin{pmatrix} u_L \\ d'_L \end{pmatrix}, \begin{pmatrix} l_L \\ \nu_L \end{pmatrix} \tag{1.26}$$

$$\phi_R = u_R, d'_R, l_R \tag{1.27}$$

The CKM matrix is required to be unitary which reduces the total number of independent

parameters to three real angles and one complex phase angle for the case of three quark generations. The complex phase angle is a necessary component for CP violation. An extremely useful representation of the CKM matrix is the Wolfenstein parameterization, which expands the matrix in terms of a small parameter λ: [6]

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \qquad (1.28)$$

One of the six possible unitarity requirements is

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \qquad (1.29)$$

With:

$$V_{ud}V_{ub}^* = A\lambda^3(\rho + i\eta) \qquad (1.30)$$

$$V_{cd}V_{cb}^* = -A\lambda^3 \qquad (1.31)$$

$$V_{td}V_{tb}^* = A\lambda^3(1 - \rho - i\eta) \qquad (1.32)$$

All six equation can be depicted as a closed triangle in the complex plane, but this particular triangle is special since all three sides are of similar dimensions (all sides are $\approx \lambda^3$). The other triangles have sides which go as $(\lambda, \lambda, \lambda^5)$ or $(\lambda^2, \lambda^2, \lambda^4)$. These triangles have the undesirable trait that even very precise measurement could have errors larger than the angle being measured (opposite the shortest triangle leg).

The interior angles are labelled α, β and γ as shown in figure 1.1. Non-zero values for β and γ are clear indications of a non-zero complex phase angle which is required for CP violation (zero angles $\rightarrow$ no complex phase angle $\rightarrow$ no CP violation).
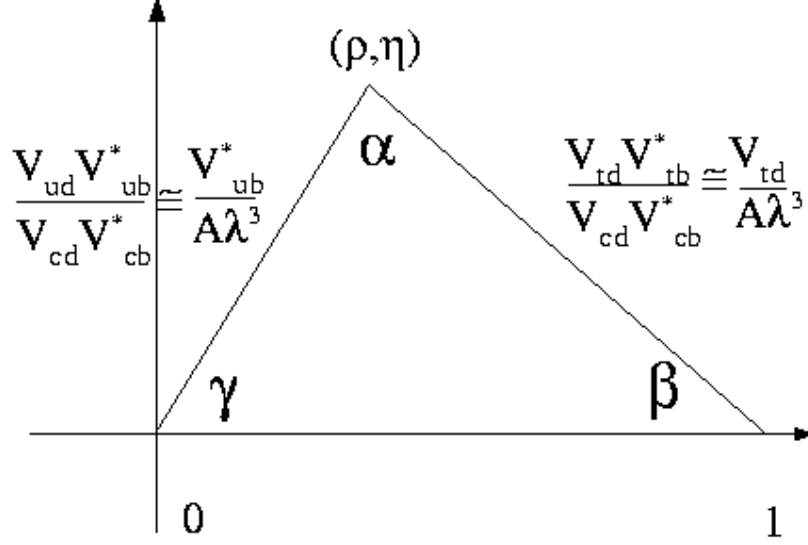
Figure 1.1: Unitarity Triangle, normalized in the $\rho - \eta$ plane.

1.4 Mixing in the B system

The mixing of B^0s can be described as a two state quantum system. The B^0s are produced in definite flavor states by strong and electromagnetic interactions, the Hamiltonian is given by $\mathcal{H}_0 = \mathcal{H}_{strong} + \mathcal{H}_{E\&M}$. The neutral Bs have two eigenstates in the flavor basis, B^0 and $\overline{B}^0$. The decay of the neutral Bs is governed by the total Hamiltonian, $\mathcal{H} = \mathcal{H}_{strong} + \mathcal{H}_{E\&M} + \mathcal{H}_{Weak}$. The eigenstates of the total Hamiltonian are the mass eigenstates, B_H and B_L, where the subscripts denote "heavy" and "light". The mass eigenstates have different time evolutions, which means the flavor of the neutral B also evolves in time. Thus a B^0 produced in a flavor eigenstate will oscillate between a B^0 and $\overline{B}^0$ until it decays.

Mixing between B^0 and $\overline{B}^0$ doesn't require CP violation, but it does require common states to which both particles can couple. Both particles can decay into $D\overline{D}$, or any state composed of c, d, $\overline{c}$, $\overline{d}$, as shown in figure 1.2. Mixing is also generated from
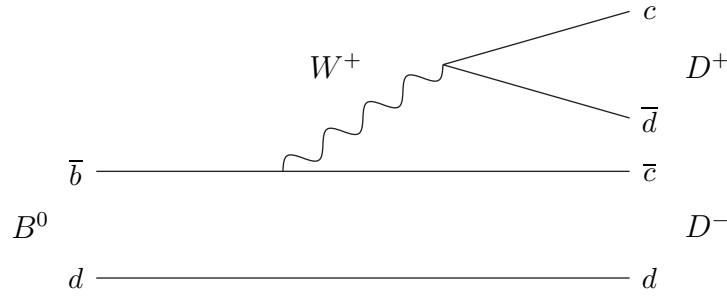
Figure 1.2: This is $B^0 \to D^+D^-$. The state D^+D^- is common to both B^0 and $\overline{B}^0$.

virtual transitions, as shown in figure 1.3. The mass splitting is dominated by the virtual transitions while the real transitions contribute to the lifetime difference [9]. The virtual transitions are suppressed by the GIM mechanism. The GIM mechanism cancellations aren't exact due to the quark mass differences in the propagators. Once the propagators are expanded out and the leading terms subtracted off, the process is dominated by a term with the top quark mass [5]. This happens because each of the three quarks type terms (u,c,t) are suppressed by the same Cabbibo angle factor, λ^3, thus top quark's large mass becomes the deciding factor. This situation is quite different from mixing in the K^0 system, where the top quark vertex is highly suppressed compared to the u and c quarks. In the K^0 system, the charm quark term dominates the virtual transitions.

A generically mixed neutral B can be written as a linear combination of flavor eigenstates [7] [8]

$$a|B^0> + b|\overline{B}^0>$$ (1.33)

this state will evolve according to the Schroedinger equation

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = H\begin{pmatrix} a \\ b \end{pmatrix} \equiv (M - \frac{i}{2}\Gamma)\begin{pmatrix} a \\ b \end{pmatrix}$$ (1.34)

The components of the Hamiltonian, M and Γ are 2×2 Hermitian matrices. The
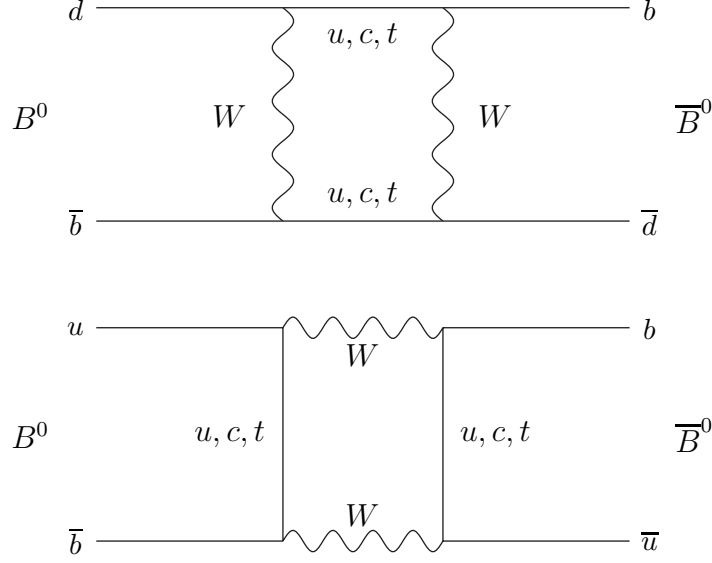
Figure 1.3: Box diagrams: exchange (top), annihilation (bottom)

two diagonal terms of $\mathcal{H}$ are equal because of CPT (particles and anti-particles must have the same mass, lifetime, etc...) while the diagonal terms of M and Γ are complex conjugates (because they are Hermitian). This means the Schroedinger equation can be written as:

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \overline{m} - \frac{i}{2}\overline{\gamma} & \delta m - \frac{i}{2}\delta\gamma \\ \delta m^* - \frac{i}{2}\delta\gamma^* & \overline{m} - \frac{i}{2}\overline{\gamma} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \tag{1.35}$$

The eigenvalues can be solved for exactly:

$$\mu_\pm = \overline{m} - \frac{i}{2}\overline{\gamma} \pm \sqrt{|\delta m|^2 - \frac{i}{2}\delta m\delta\gamma^* - \frac{i}{2}\delta m^*\delta\gamma - \frac{1}{4}|\delta\gamma|^2} \tag{1.36}$$

It is expected that $\Delta\Gamma_{B_d} \ll \Gamma_{B_d}$, since the decay channels common to both B^0 and $\overline{B}^0$ currently have branching fractions measured to be at or below 10^{-3}. Since the channels contribute with differing signs, the combined total should still be on the order of 10^{-3} [7]. The ratio

$$x_d = \Delta m_{B_d}/\Gamma_{B_d} = 0.730 \pm 0.029 \text{ [6]} \tag{1.37}$$

shows that Δm_{B_d} and Γ_{B_d} are of the same order, thus $\Delta m_{B_d} \gg \Delta\Gamma_{B_d}$. The $\delta\gamma$ term can be safely neglected, at which point the eigenvalues can now be written as:

$$\mu_H = \overline{m} - \frac{i}{2}\overline{\gamma} + |\delta m| \tag{1.38}$$

$$\mu_L = \overline{m} - \frac{i}{2}\overline{\gamma} - |\delta m| \tag{1.39}$$

The mass splitting of the two states is $\Delta m = 2|\delta m|$. The eigenstates are:

$$|B_H\rangle = \frac{1}{\sqrt{2}}\left(|B^0\rangle + \frac{|\delta m|}{\delta m}|\overline{B}^0\rangle\right) \tag{1.40}$$

$$|B_L\rangle = \frac{1}{\sqrt{2}}\left(|B^0\rangle - \frac{|\delta m|}{\delta m}|\overline{B}^0\rangle\right) \tag{1.41}$$

The time dependence of the physical states can be written:

$$|B^0_{phys}(t)\rangle = e^{-i(\overline{m} - \frac{i}{2}\overline{\gamma})t}\left[\cos(\frac{\Delta mt}{2})|B^0\rangle - i\frac{\delta m}{|\delta m|}\sin(\frac{\Delta mt}{2})|\overline{B}^0\rangle\right] \tag{1.42}$$

$$|\overline{B}^0_{phys}(t)\rangle = e^{-i(\overline{m} - \frac{i}{2}\overline{\gamma})t}\left[\cos(\frac{\Delta mt}{2})|\overline{B}^0\rangle - i\frac{\delta m}{|\delta m|}\sin(\frac{\Delta mt}{2})|B^0\rangle\right] \tag{1.43}$$

1.5 Application to sin(2β) and the CKM matrix

The measurement of Δm_{B_d} has several applications for other physics results. The semileptonic data set is much larger than fully reconstructed B^0 decay modes to CP states, thus the sin(2β) measurement benefits from several precisely measured parameters (oscillation frequency, mistag rates) which otherwise would affect the statistical and systematic sensitivity of the sin(2β) measurement.

The Δm_{B_d} measurement also helps to put a constraint on the unitarity triangle (figure 1.1). The leading term for the virtual $t\overline{t}$ contribution in the box diagram (figure 1.3) gives an equation for the mass difference in terms of the CKM parameters V_{tb} and V_{td}:

$$\Delta m_{B_d} = 2\delta m = \frac{G_F^2}{6\pi^2}B_B f_B^2 m_{B_d} m_t^2 \eta f(x_t)|V_{tb}^* V_{td}|^2 \tag{1.44}$$

where B_B is the bag parameter used to characterize the hadronic matrix elements, f_B is the B meson decay constant, η is a QCD correction, m_{B_d} is the B^0 mass, m_t is the top quark mass and $f(x_t)$ gives the electroweak loop contributions of the top quark without QCD corrections and $x_t = m_t^2/m_W^2$:

$$f(x_t) = \frac{4 - 11x_t + x_t^2}{4(1 - x_t)^2} - \frac{3x_t^2 \ln x_t}{2(1 - x_t)^3} \tag{1.45}$$

The length of the triangle side opposite the angle γ depends on $V_{tb}V_{td}$. A measurement plus uncertainties of m_{B_d} creates an annulus in the ρ, η plane which constrains where the peak of the triangle can be located, shown in figure 1.4. The largest contribution to the uncertainty in the theoretical calculation of Δm_{B_d} is the term $B_B f_B^2$. The value of $B_B f_B^2$ is obtained from lattice calculations which will improve over time. A directly measurement of these quantities is possible, however, the amount of data required to improve over current calculations is estimated to be the amount of data collected by BaBar over its entire life. Until the uncertainty on $B_B f_B^2$ is reduced, a better measurement of Δm_{B_d} doesn't really improve the errors on $V_{tb}V_{td}$.
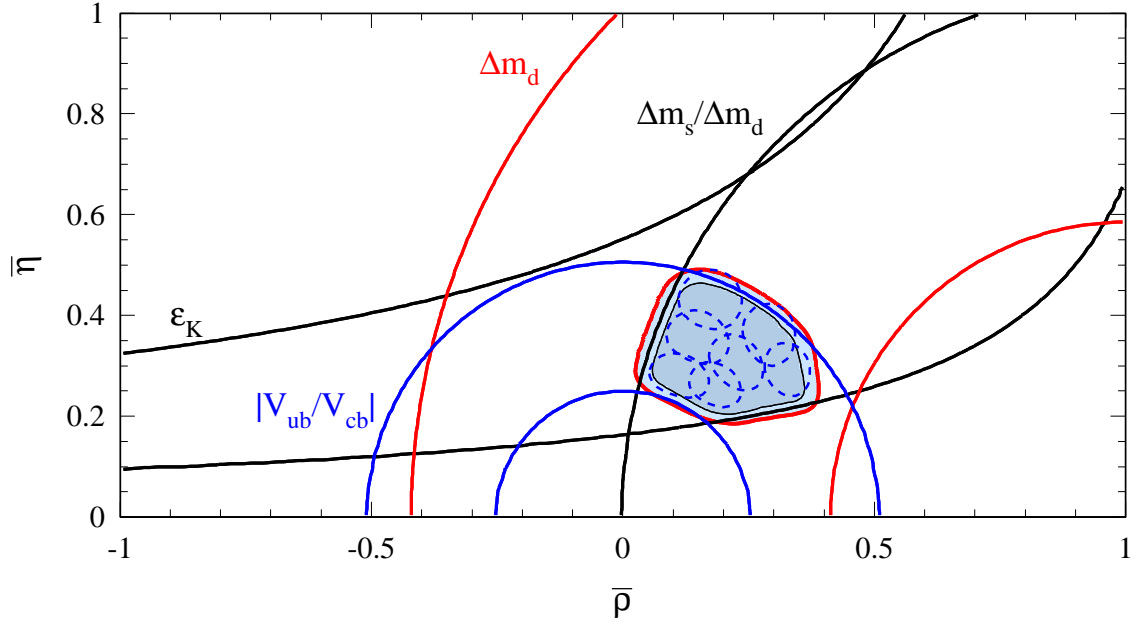
Figure 1.4: Constraints applied by Δm_{B_d} in the $\rho - \eta$ plane. The ϵ_K curves are obtained from measuring CP violation in $K^0 - \overline{K}^0$ mixing. The $|V_{ub}/V_{cb}|$ constraint comes from measurements of the decay $B \rightarrow X_u l\nu$. The Δm_{B_d} constraint comes from equation 44. The upper limit can be improved by using a ratio of $\Delta m_s/\Delta m_d$. Δm_s measurements have yielded a lower limit of $\Delta m_s > 12.4 \ ps^{-1}$ at the 95% confidence level [6].
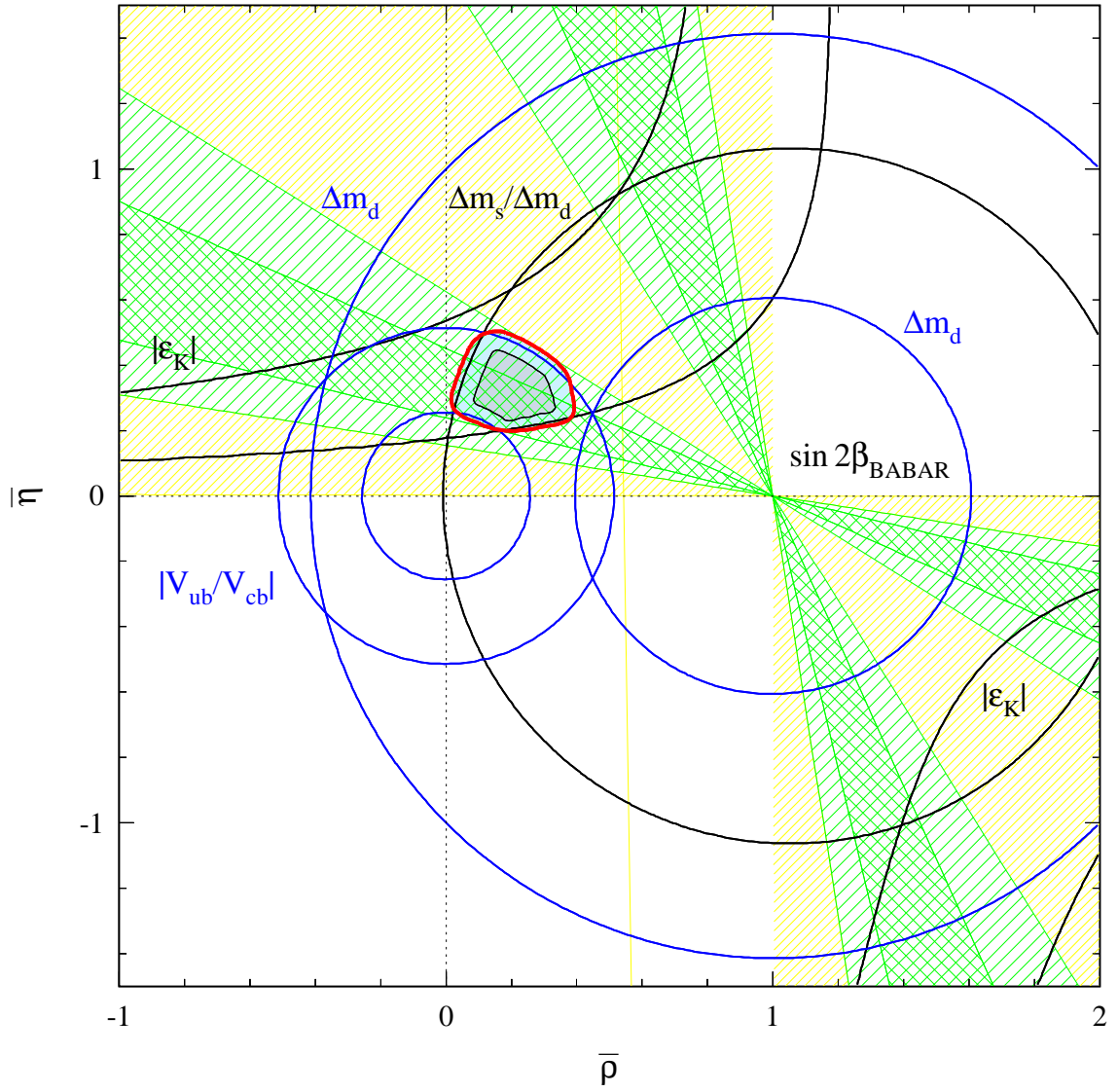
Figure 1.5: This figure shows the same constraints as the previous plot plus sin2β. The measurement does not yield a unique value for β, but the ambiguity can easily be resolved with information from the other constraints.

Chapter 2

The BaBar detector

2.1 Overview

The primary goal of the BaBar experiment is to accurately measure sides and angles of the unitarity triangle in an attempt to over-constrain the triangle. These measurements will improve our knowledge of CP violation and possibly challenge the Standard Model if the triangle sides and angles violate unitarity. BaBar is capable of performing other important physics measurements, including other B physics, τ, charm and two photon physics. Because CP asymmetries in the B system are expected to be large, measurements with an accuracy of 10% can be made with several hundred reconstructed events, however, typical branching ratios are on the order of 10^{-5}, thus tens of millions of B^0 pairs must be produced. To deliver this many B^0's to the BaBar detector, the design luminosity of PEP-II is $3\times10^{33}cm^{-2}s^{-1}$ which will deliver $\approx$ 3 B pairs per second. The BaBar detector is composed of five major concentric subdetectors,

- Silicon Vertex Tracker (SVT)

- Drift Chamber (DCH)
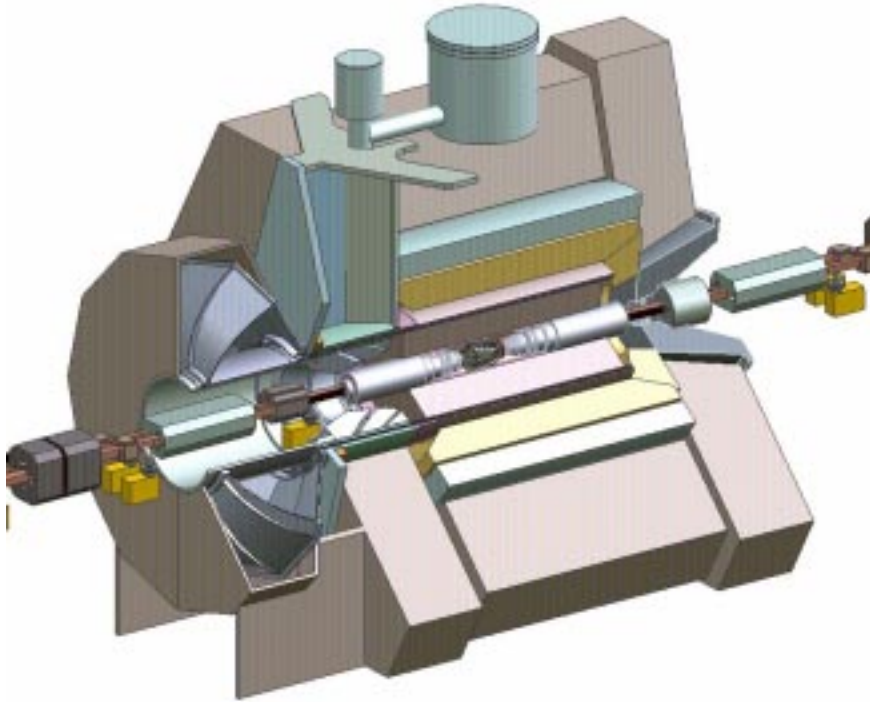
Figure 2.1: BaBar detector.

- DIRC (Detection of Internally Reflected Cherenkov Light)

- Electromagnetic Calorimeter (EMC)

- Instrumented Flux Return (IFR)

and a super-conducting solenoid located between the DIRC and EMC which provides a uniform 1.5 T magnetic field.

2.2   PEP-II

The SLAC Linac is a two mile long linear accelerator which provides high energy electrons and positrons to the PEP-II storage ring. The linac was completed in 1966 and has seen several upgrades and changes over its 36 year history. The PEP-II stor-
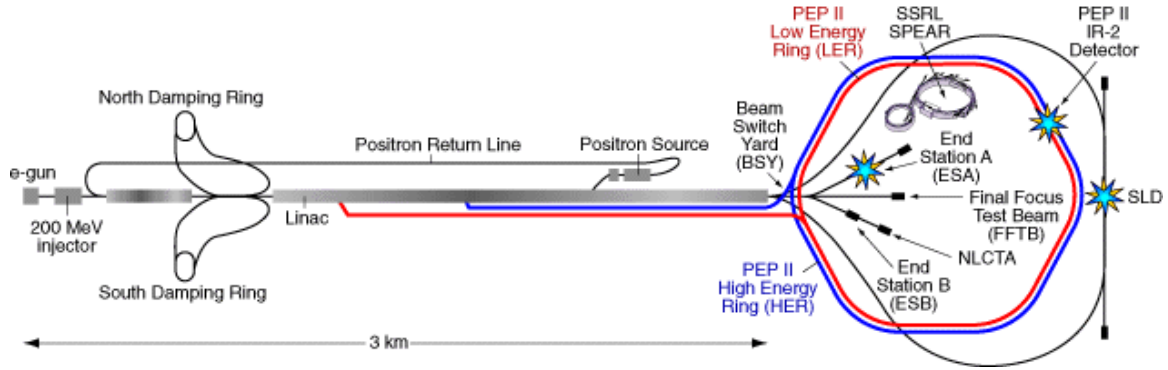
Figure 2.2: SLAC linac and storage rings.

age ring is also an upgraded machine. It started out as PEP (Positron-Electron Project), a 30 GeV storage ring used through the 1980's to test the theory of Quantum Chromodynamics (QCD). The PEP tunnel is 800 meters in diameter and currently houses the PEP-II storage ring. The most striking feature of PEP-II is that it delivers the electrons and positrons with different beam energies.

2.2.1 Asymmetric Collider

The asymmetric mode of operation is a novel solution to a very challenging problem. The problem is that when the $\Upsilon(4s)$ (mass = 10.58 GeV) decays into two B^0's (mass = 5.28 GeV each), there is very little energy left over in the form of kinetic energy. Each B^0 will travel along with a momentum of 300 MeV in the $\Upsilon(4s)$ rest frame. Thus the B^0 will only travel about 26μm before it decays. This is too small for current vertex trackers to accurately resolve. By colliding asymmetric beams, the $\Upsilon(4s)$ is given a large boost which increases both the velocity and time dilation for the decaying B^0's. In BaBar, 9.0 GeV electrons are collided with 3.1 GeV positrons to produce a $\Upsilon(4s)$ with an energy of 12.1 GeV and $\beta\gamma = 0.56$. This produces an average B^0 separation of 250μm, still

challenging, but much more manageable from a resolution point of view.

2.2.2   Accelerating particles

Accelerating electrons and positrons to their final energies is accomplished in several steps. Electrons are first produced by an electron gun - a device similar to the electron gun in a television set. A filament is heated and electrons are released into an area with a strong electric field. This accelerates the electrons up to 10 MeV. The beam is relatively spread out, so the electron beam is sent to the north damping ring (positrons to the south). A narrowly focused beam of particles will produce more collisions than a diffuse beam, so the damping ring's job is to dissipate motion not in the beam direction. Each time an electron completes a cycle in the damping ring it loses energy to synchrotron radiation and receives a boost equal to the synchrotron loss in the beam direction. Synchrotron radiation dissipates energy in the direction that the particle is moving, so particles with momentum transverse to the beam direction will dissipate some of this transverse momentum. The particle will then receive a boost in the beam direction, the net effect is a focusing of the beam in the beam direction.

The next step is to return the electrons to the linac where they will be accelerated to their full energy. Positrons are produced by removing off some high energy electrons from the linac and colliding them with a fixed tungsten target which produce electron-positron pairs. The positrons are collected and returned to the start of the linac where they follow the same path as electrons to the damping rings. The energy the particles gain in the linac is proportional to how far they travel in the linac. The low energy positrons are removed first, then the higher energy electrons (as shown in figure 2.2). Both beams are sent to the PEP-II storage ring, where they will ultimately be steered

together to collide in the BaBar detector.

2.3 SVT
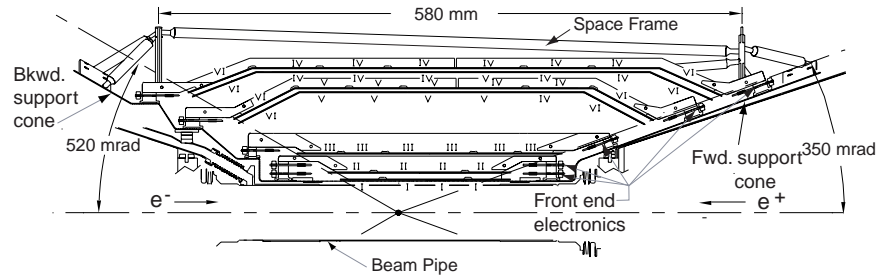
Figure 2.3: 5 layer SVT profile.

The inner most subdetector in figure 2.1 is the Silicon Vertex Tracker (SVT). The SVT has 5 double sided silicon strip detector layers, as shown in figure 2.3. The inner side has strips oriented perpendicular to the beam axis to measure z position while the outer surface has orthogonal strips to measure the ϕ coordinate. The SVT geometry was determined by a need to accommodate the final beam bending magnets, located at ± 20 cm, and cover the largest amount of solid angle as possible. The first three layers are flat modules which are critical for accurately measuring the B^0 vertices. The fourth and fifth layers have an arch design which maximizes solid angle coverage while avoiding large track incidence angles. They help the inner layers determine B^0 vertices, serve to align SVT tracks with DCH tracks and are important for tracking charged particles which have a transverse momentum less than 100 MeV, since these particles will not enter the Drift Chamber. Particular attention is paid to the radiation hardness of SVT components (the SVT will accumulate a very large radiation does due to its proximity to the beam

pipe) and also to minimizing the amount of material in the SVT to minimize the effect of multiple scattering for the other subdetectors.

Because of the beam energy asymmetry, it is extremely important to maximize coverage in the forward direction, since the asymmetric boost will increase the number of decay tracks that will travel that direction. For this reason, all instrumentation readout and cooling components are located in the backward support cone. This can be seen by the asymmetric between the shape of the forward and backward support cones in figure 2.3.

Layer	Radius (cm)	Modules	ϕ pitch (μm)	z pitch (μm)
1	3.2	6	50 or 100	100
2	4.0	6	55 or 110	100
3	5.4	6	55 or 110	100
4a	12.4	8	100	210
4b	12.7	8	100	210
5a	14.0	9	100	210
5b	14.4	9	100	210

Table 2.1: SVT layers, radius and pitch.

The SVT is the single most important subdetector for a mixing measurement. The SVT resolution completely dominates the time difference determination while all subdetectors play varying roles in the mixing status determination. The SVT must be able to discriminate between the two vertex positions which have an average separation distance of 250 μm. Figure 2.4 shows the resolution in both z and ϕ as a function of incident track angle. The SVT resolution is very good as far as the mixing measurement is concerned. The SVT resolution is dominated by multiple scattering, not SVT hardware.
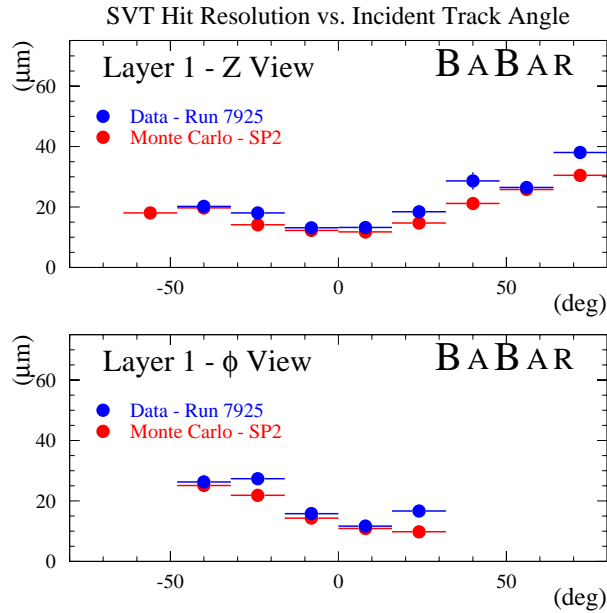
Figure 2.4: SVT resolution for data and MC as a function of incident angle.

2.4 DCH

The Drift Chamber(DCH) is the main tracking device for BaBar. It provides a precision measurement of transverse momentum (p_T) for charged tracks with p_T greater than 120 MeV (tracks below this threshold generally don't have enough DCH hits to form good tracks). The DCH starts at a radius of 23.6 cm and extends out to 79.0 cm. The DCH contains 7104 cells, each approximately 1.2cm by 1.8cm. The 40 layers are organized into 10 superlayers, each four wire superlayer has uniform stereo angle alignment inside of a superlayer. Four of the superlayers (1,4,7 and 10) are axial (A) superlayers - the wires run perfectly parallel to the detector's z axis. The remaining superlayers have a non-zero stereo angle which allows the DCH to determine a tracks z position as well as radius and ϕ. The stereo angle ranges from 45 mrad in superlayer 2 out to 76 mrad in super-layer 9

and the sign of the angle alternates every other stereo layer creating two different types

of stereo superlayers (U - positive stereo angle, V - negative stereo angle).
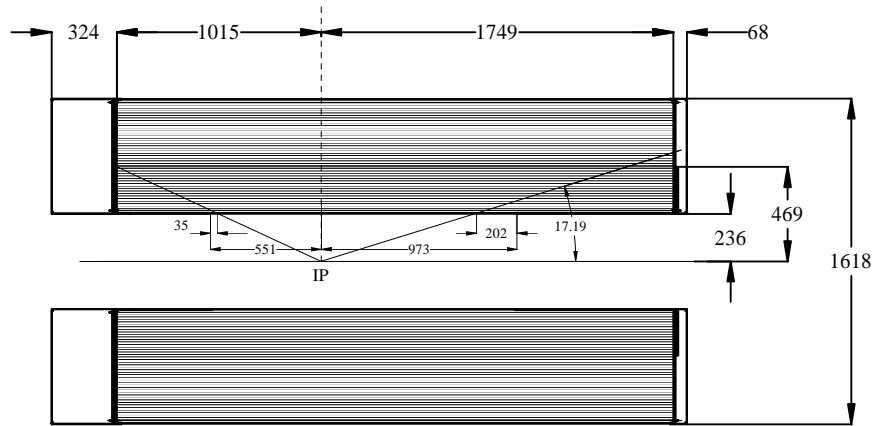
Figure 2.5: Side view of the Drift Chamber layout, the high energy electrons are traveling to the right. The Interaction point (IP) is located closer to the high energy side to improve solid angle coverage.

Superlayer	Cells/layer	Inner radius (cm)	Layer Type
1	96	26.0	A
2	112	31.9	U
3	128	37.1	V
4	144	42.3	A
5	176	48.1	U
6	192	53.3	V
7	208	58.5	A
8	224	64.3	U
9	240	69.5	V
10	256	74.7	A

Table 2.2: DCH superlayers, layer type is axial(A) or stereo(U,V).

Each cell consists of a single sensing wire, maintained at a voltage of 1900V - 1960V, and six field shaping wire. For cells on the inner side of a superlayer, the field shaping wires are grounded on on the rear end plate. For cells on the outer side, one wire is held at 350V. The resolution for each cell is shown in figure 2.6, 125 μm is the average resolution.
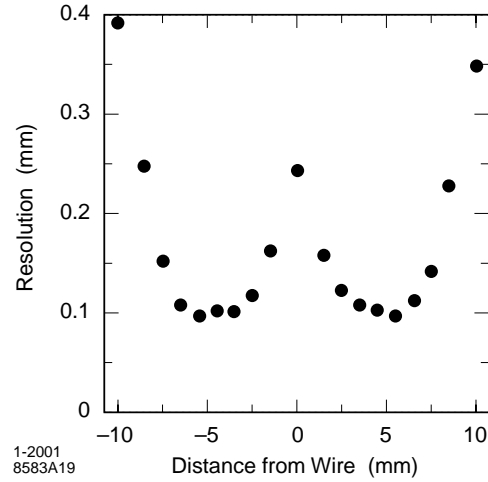
Figure 2.6: Drift Chamber resolution vs track distance from wire at 1961V.

The DCH uses a mixture of He:ISO-butane (80%:20%) as its counting gas. This mixture has a low density (minimize multiple scattering for good p_T resolution) and good spatial and dE/dx resolution and reasonably short drift time. The gas and wires total to 0.3% radiation lengths (X_0). The inner cylinder (1mm beryllium) contributes 0.28% X_0 while the outer carbon fiber cylinder contributes 1.5% X_0.

The DCH serves several purposes other than tracking. The DCH provides prompt information to the Trigger system which is used as a course filter for removing unwanted events and saving valuable processing time. The DCH also has a crucial role in particle identification (PID). Figure 2.7 shows how a particle's energy loss per unit distance (dE/dx) vs. momentum correlates to a particular particle. This information can be combined with information from other subdetectors to estimate the PID of charged tracks in each event.
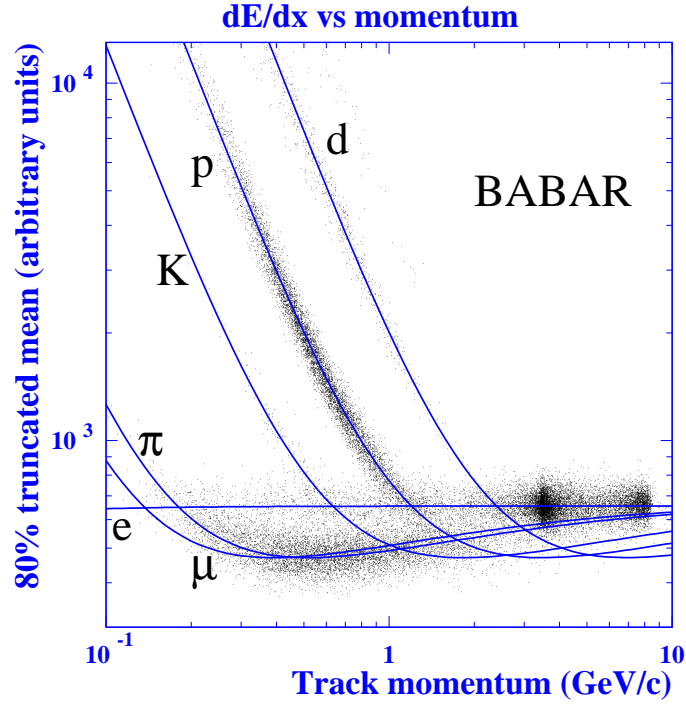
Figure 2.7: dE/dx vs p from Drift Chamber.

2.5 DIRC

The DIRC consists of 144 quartz bars arranged into 12 "bar boxes". Each bar is 1.7 cm thick, 3.5 cm wide and 490 cm long. Charged particles traveling through a bar will radiate Cherenkov radiation whenever the particle is traveling above the speed of light in quartz (= c/(index of refraction)). The Cherenkov light is guided to the instrumented section of the DIRC via total internal reflections with the quartz bar faces. The uninstrumented end is mirrored to return photons back to the instrumented end.

The instrumented end is a large, half toroidal tank filled with purified water which closely matches the quartz bar's index of refraction to minimize transmission loses. The angle between the light cone and the emitting particle is preserved by the internal reflections. Once the light gets to the water tank, the photons travel through the
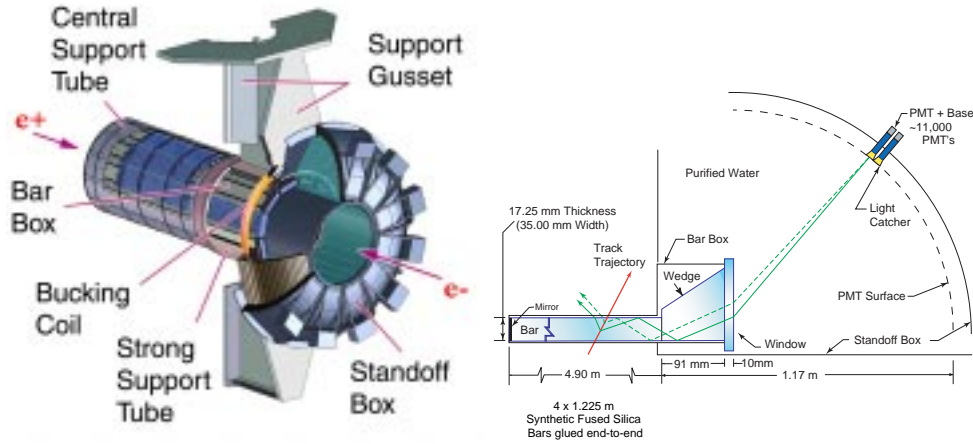
Figure 2.8: Left: DIRC quartz bars and standoff tank. Right: DIRC operating principle

water until detected by one of the 11,000 Photomultiplier Tubes (PMTs). This princi-ple is shown in figure 2.8. The large standoff box radius allows for a finer resolution of Cherenkov angle. The resolution for a single photon is 10.2 mr, as shown in figure 2.9. There are generally around 30 or so photons detected each event so by using pattern recognition, the DIRC can improve angular resolution down to about 2.8 mr. This resolu-tion allows for a 3 sigma separation between pions and kaons at about 3 GeV. The DIRC is extremely valuable for particle identification.
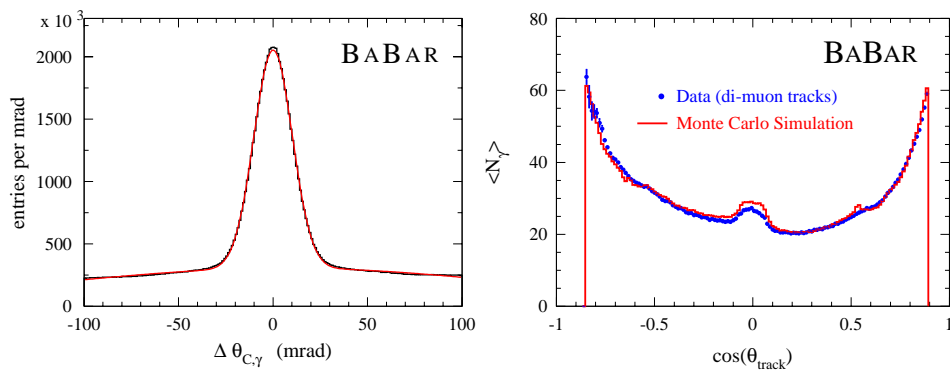
Figure 2.9: DIRC single photon resolution and photons per track.

2.6   EMC

The electromagnetic calorimeter is composed of 6580 CsI crystals. 5760 of these crystals are contained in the barrel region divided in to 48 polar-angle rows (θ direction) with 120 crystals in each row (ϕ direction). The forward end cap region contains the remaining crystals in eight polar-angle rows. Each crystal a trapezoidal pyramid with a front face surface area of 5cm$\times$5cm and a depth of 29.76-32.55 cm, the crystal length increases as you move in the forward direction (which corresponds to 16-17.5 radiation lengths). All cooling, electronics and almost all support material is located behind the crystals to minimize the material in front of the crystals.

The calorimeter works by absorbing photons or interacting electromagnetically with charged particles from an event and subsequently emitting scintillation light. The scintillation light is detected by two photodiodes on the back of the crystals and turned into digital signals. The photon energy resolution and angular resolution at 90 degrees is shown in figure 2.10.
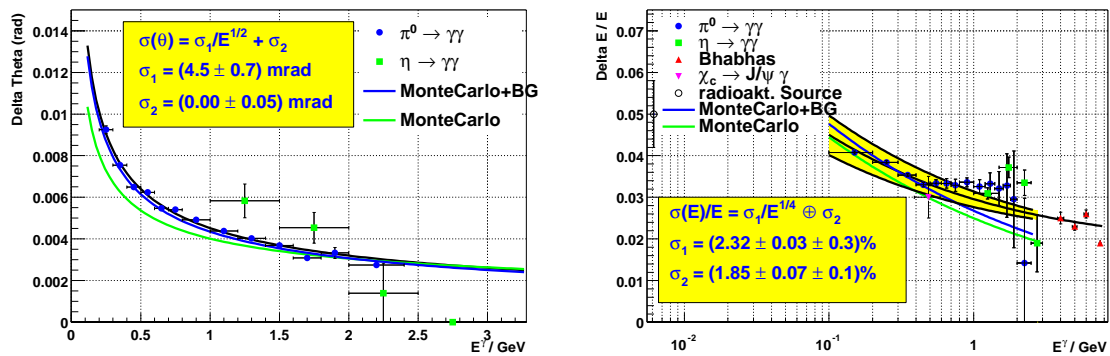
Figure 2.10: EMC position and energy resolution.

An important issue for the calorimeter is calibration. The energy detection range

extends from 10 MeV up to 9 GeV and is calibrated in four separate, overlapping regions. Several methods of calibration are available:

- Injection of charge: Charge is injected at the preamplifier input, with predetermined magnitude, phase and pattern. This tests in detail the response of each amplifier. This calibration takes several minutes.

- Bhabha events: e^+e^- pairs can be used for calibration, running time for calibration is less than one day of running. This calibration data is taken simultaneously with normal data taking.

- Liquid radioactive source: 6.1 MeV photons from ^{16}N decays.

- Light pulser: Convenient for tracking short term changes in crystal response.

Calibration helps track short and long term changes in the crystal response. Crystal response can vary due to radiation damage or damage to optical surfaces and couplings. Calibration also measures constants used for maximizing the calorimeters resolution.

2.7 IFR

The Instrumented Flux return serves both as a flux return for the 1.5 T solenoidal magnet and as a muon and neutral hadron detector. The IFR has 3 main segments: the barrel, forward and backward end-caps. The barrel region extends from a radius of 1.88 m to 3.23 m and is divided into sextents. The end-caps are hexagonal plates divided into two halves to allow the IFR to be moved into place around the inner subdetectors. The IFR consists of 18 iron plates which vary in thickness. Detailed Monte Carlo studies

showed that low momentum muon and K_L^0 detection improve with thinner absorbing plates, but this effect is only significant for the first absorption length. As a result, the first nine layers of the IFR are 2 cm thick, then 4 at 3 cm, followed by 3 at 5 cm and 2 at 10 cm. In the end caps, one of the 10 cm plates is replaced with a 5 cm plate.

The IFR is instrumented with Resistive Plate Chambers (RPCs). An RPC consists of two 2 mm thick Bakelite plates separated by 2 mm, with outer surfaces painted with graphite of high surface resistivity (100kΩ/square) and covered with an insulating film. One graphite surface is connected to high voltage (8kV) while the other is connected to ground. The inner surfaces (the two Bakelite surfaces which face each other) are treated with linseed oil to enhance high efficiency and low noise. Each RPC is 3.2 cm tall, 125 cm wide and 181 cm to 320 cm long (depending on location of RPC). The active space between the two plates is filled with a gas mixture of Argon, Freon and a very small amount of ISO-butane. This gas mixture is non-flamible and environmentally safe.

The RPC operates when a charged particle crosses the chamber producing a quenched spark which produces a signal. There are 21 active detector layers. A double layer cylindrical RPC resides between the EMC and IFR, 17 layers exist between the 18 IFR iron plates and two more active layers are located just outside the last IFR plate. There are 18 active layers in the end caps (one prior to the first IFR plate plus 17 in between plates).

2.8   Trigger

The BaBar Trigger is composed of the Level 1 Trigger (L1 Trigger) and the Level 3 Trigger (L3 Trigger). The Trigger systems job is to quickly filter out uninteresting events

while maintaining 100% efficiency for B physics events.

2.8.1 L1 Trigger

The L1 Trigger is a hardware based filtering system. The L1 Trigger is comprised of the Drift Chamber Trigger (DCT), Electromagnetic Calorimeter Trigger (EMT), Instrumented Flux Trigger (IFT) and the Global Level 1 Trigger (GLT). The DCT produces three 16-bit maps of ϕ coordinates for candidate tracks which pass selection as either a short track (p_T >120 MeV track which reaches Super Layer 5), a long track (p_T >180 MeV track which reaches Super Layer 10) or a high momentum track (p_T >800 MeV). The EMT produces five 20-bit ϕ maps of particle candidates depending on the amount of energy deposited and location, see table 2.3. These maps plus a single ϕ map from the IFT are collected by the GLT to form 24 trigger lines which are passed to the L3 trigger, see figure 2.11.

EMT objects			DCT objects		
	Cluster Description	Energy Cut		Description	P_t Cut
M	Minimum ionizing	100 MeV	B	Track reaching SL-5	120 MeV
G	Intermediate ionizing	300 MeV	A	Track reaching SL-10	180 MeV
E	High energy e/γ	700 MeV	A'	High p_t track	800 MeV
X	forward endcap	100 MeV		reaching SL-10	
Y	Backward barrel	1 GeV			

Table 2.3: Trigger ϕ map objects defined for the EMT and DCT. The IFT has a single ϕ map labelled U which indicates a track in the IFR.

The L1 Trigger system was designed to have two highly efficient orthogonal triggers, the DCT and EMT. Most of the 24 Trigger lines (13/24) are purely DCT or EMT triggers. This allows the trigger efficiencies to be easily cross checked and measured. For $B^0\overline{B}^0$ events, each trigger is > 99% efficient alone and > 99.9% combined.

After the GLT collects all the L1 Trigger information, it decides whether to issue
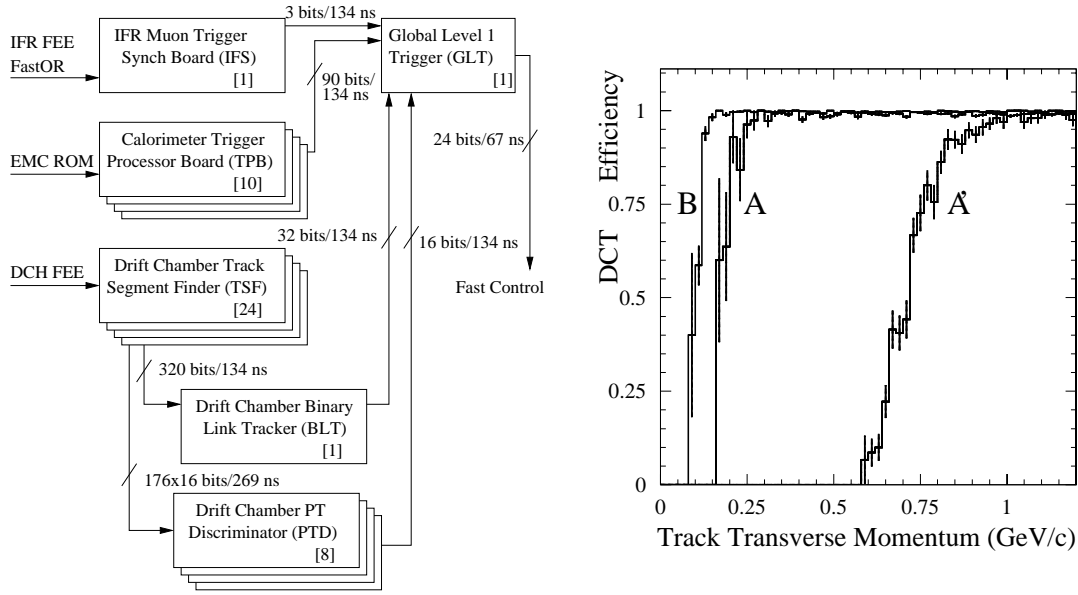
Figure 2.11: Left: Elements which make up the Level 1 trigger. The TSF sends track segment information to the BLT and PTD. The BLT, PTD, EMT and IFT all send their respective bit map information to the GLT which decides if a Level 1 accept should be issued. Right: DCT efficiency plot. The A and B tracks turn on curves are determined by the minimum p_t required to reach half way and all the way through the Drift Chamber. The A' track turn on is programmable, currently set at 800 MeV.

a L1 Trigger accept or not. If a L1 accept signal is sent by the GLT to the other trigger boards, the event data currently stored in memory is passed to the L3 Trigger, if no L1 accept is received, the data is overwritten and lost. A L1 accept decision sends L1 information to L3 within 11-12 μs after the corresponding event occurred.

2.8.2 L3 Trigger

The L3 Trigger is a software based filter which takes the greatly reduced number of events from L1 Trigger and applies further constraints on selecting good events to be stored on disk for future use. The L3 trigger runs in parallel on 32 computer nodes. L3 takes data from the DCH and EMC and forms track and cluster objects. These objects are passed through a series of filters designed to eliminate backgrounds such as tracks not

coming from the interaction region. Events which pass the L3 filters are written to disk

to be processed later. Figure 2.12 shows the L3 event display. This display shows the raw

tracks (red lines), EMC clusters (red towers) and activated trigger lines (2E, EM*, etc...).
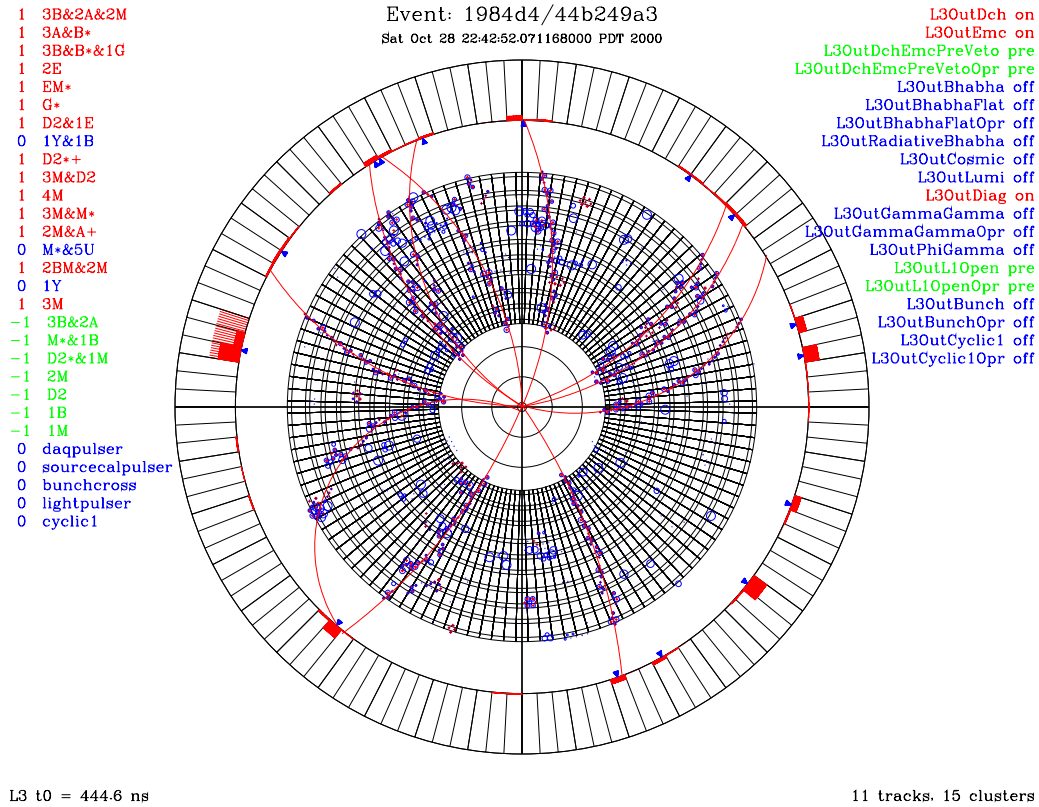
Figure 2.12: L3 event display. The blue circles are track segments from the TSF boards. The red lines label reconstructed charged tracks and the red towers show energy deposits reported by the EMT.

The L3 trigger creates composite objects from the basics EMT and DCT objects. The trigger line definitions change over time, figure 2.12 shows trigger lines used on Oct 28, 2000. The naming conventions used by the L3 trigger are shown in table 2.4.

Back-to-Back objects		
Object	Description	ϕ Cut
B^*,A^*	back to back short, long tracks	124°
M^*,G^*	back to back M, G clusters	117°
E-M	E and M clusters back to back	126°
DCT+EMT matching objects		
BM	B and M with matching phi	$< 27^o$
AM	A and M with matching phi	$< 27^o$
A'M	A' and M with matching phi	$< 27^o$
Compound Trigger objects		
A+	1A & 1A'	
D2	2B & 1A	
$D2^*$	B^* & 1A	
$D2^*+$	B^* & 1A+	
Z^*	2Z with loose back-to-back cut, Z is any primitive	

Table 2.4: Trigger objects defined for the EMT and DCT. The IFT has a single object labelled U which indicates one or more tracks in the IFR.

2.8.3 Trigger Upgrades

The L1 and L3 triggers have lived up to all design expectations, but as BaBar pushes up its luminosity the trigger rates will approach design limits and create problems for the Data Acquisition system (DAQ). There are very few ways to improve the current trigger rates with the existing hardware. One way is to drop some of the higher rate pure DCT and EMT trigger lines, replacing them with composite lines which require both DCT and EMT objects. The L3 trigger group maintains the trigger line definitions and can adjust them if needed, however, there is a strong desire to maintain the two orthogonal triggers as much as possible.

Another way to improve the trigger rates is with hardware upgrades. The DCT group currently has an improved version of the PTD which not only cuts on a charged track's p_t, but can determine if the arc of the track roughly intercept the interaction point (IP). This is effectively a cut on the point of closest approach (POCA) of the track and can

eliminate some tracks which originate too far from the IP.

This new PTD hasn't been used yet, but rather it is being expanded further into a whole new set of trigger boards which will ultimately be able to cut on track POCA and z position. The ability to cut in z is extremely important, figure 2.13 shows the distribution of z positions for tracks passing L3. The two bumps at $\approx \pm 20$ cm are backgrounds caused by the final bending magnets. The ability to cut at ± 10 cm would greatly improve the trigger rates.

The details of where and how effective the z cut will be is still being worked out, but even a loose cut out at ± 20 or 30 cm will still be a strong improvement over the current design. One complication is that this z cut is derived from DCT objects, there is currently no prospects for improving the EMT rates. The full benefit of the z cut requires that a DCT object be combined with all EMT objects. This is probably inevitable, but the L3 group will try to maintain the orthogonal triggers as long as possible
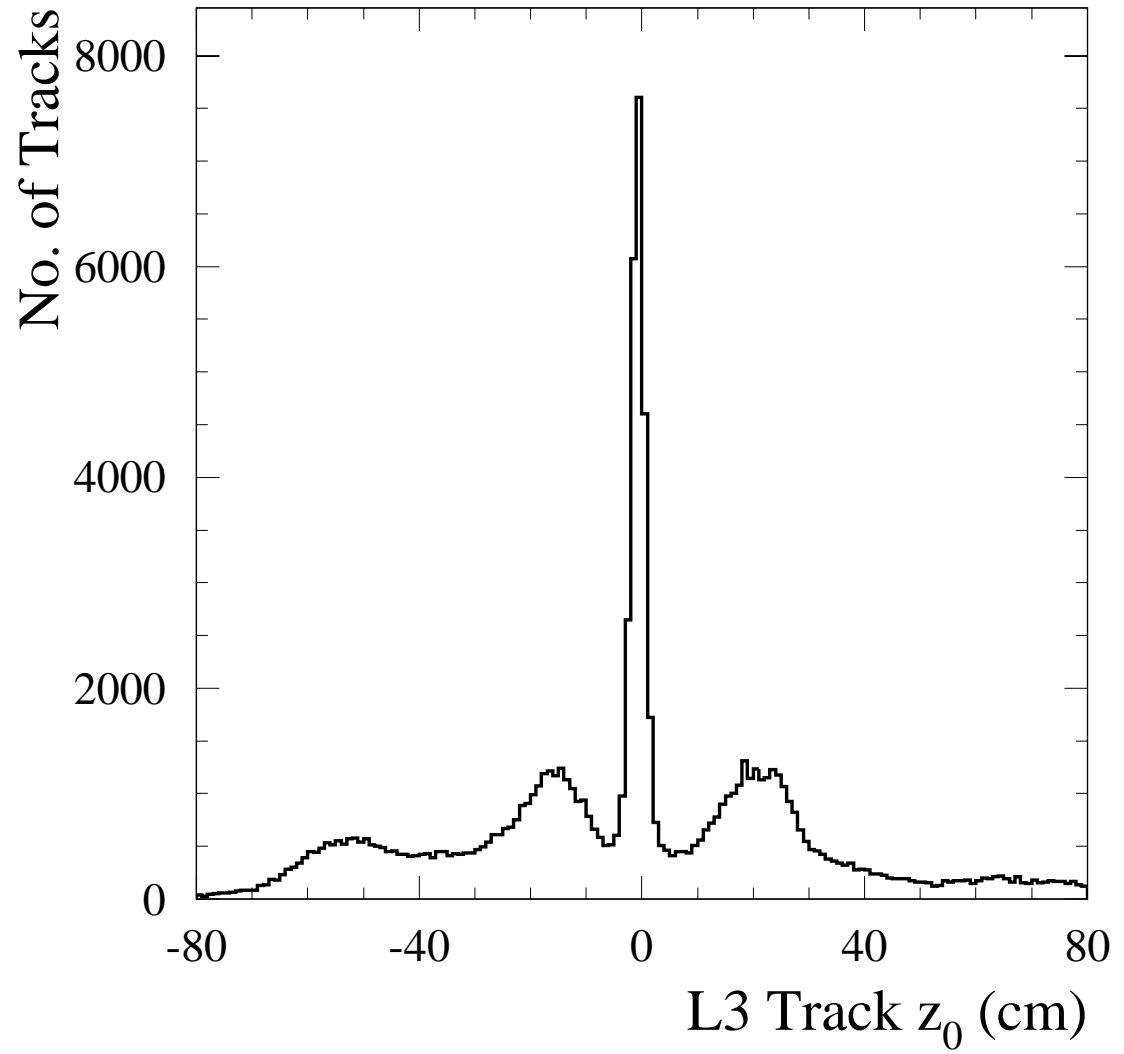
Figure 2.13: Z position distribution of tracks which pass L3.

Chapter 3

Analysis Overview

This section is a brief overview of the analysis and helps to motivate and map out the next four chapters. In a perfect world, one would simply take a pure sample of $D^*l\nu$ events, record the time difference (Δt) and mixing status of each event and fit this distribution with the theoretical expectation given in section 1.4 to determine Δm_d. This basic recipe serves as the framework for this measurement, but each step requires a considerable amount of work to actually accomplish:

1. **Pure sample of $D^*l\nu$ events**: Unfortunately, there are no pure samples of $D^*l\nu$ events. As a result, extensive work must be done to create a sample of $D^*l\nu$ events as pure as possible and completely characterize the nature of the remaining backgrounds. This is covered by chapter 4.

2. **Record the time difference**: The time difference measurement is complicated by detector resolution effects, lack of knowledge about the $\Upsilon(4s)$ decay vertex and certain approximations which are required due to the missing neutrino. The determination of all the errors and biases associated with the Δt measurement is the most difficult aspect of the Δm_d measurement. This is covered by chapter 5.

3. **Record the mixing status**: The accuracy of the mixing status determination depends heavily on the topology and particle types resulting from the non-reconstructed B^0 decay. It is important to be able to determine the mixing status and the associated error rate (mistag rate). This is covered by chapter 6.

4. **Fit the result to theoretical expectations**: The data sample is not pure, the Δt distribution is smeared, the mistag rate varies and the final fit must account for each of these issues. A very involved fitting procedure and various background samples are needed to disentangle the effects of backgrounds, mistag rate and resolution on the mixing frequency measurement. This is covered by chapter 7.

The final result of this prescription is a measurement of Δm_d, the B^0 lifetime, mistag rates and resolution parameters plus all associated errors, statistical and systematic.

Chapter 4

Event Selection

This analysis determines Δm_d by using only $B \to D^*l\nu$ decays. The $D^*l\nu$ channel's major attraction is its large branching fraction. Because the neutrinos escape detection and measurement, we are missing a powerful constraint that the hadronic B^0 decay modes can utilize to ensure all the decay daughters add up to a B^0. As a result, the $D^*l\nu$ channel requires very careful analysis of backgrounds and errors.

The full decay channels considered in this analysis are $B \to D^*l\nu$, $D^* \to D^0\pi$ and the D^0 can decay into four different final states: $K^-\pi^+$, $K\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$, $K_S\pi^+\pi^-$. The $K^-\pi^+$ mode is by far the cleanest mode, the $K\pi^+\pi^0$ and $K^-\pi^+\pi^+\pi^-$ modes have larger branching ratios but larger backgrounds and the $K_S\pi^+\pi^-$ mode is fairly clean, but harder to reconstruct due to the neutral K_S. The $B \to D^*l\nu$ is interesting since its great statistics provide an early measurement of Δm_d and it also provides a useful cross check for the measurement using hadronic modes only.

4.1   Data Set

The data used for this analysis comes from "good runs" starting in January 1999, ending in October 2000. The set consists of $20.6 fb^{-1}$(on-peak) + $2.6 fb^{-1}$(off-peak). Table 4.2 gives a detailed list of runs used for this analysis.

A significant amount of Monte Carlo data was processed and used for validation studies. The number of events and equivalent luminosity are summarized in Table 4.1 [25].

Mode	Events	Equv.Lumi (fb^{-1})
B^0	7968 k	14.49
B^+	9288 k	16.88
Filtered B^0	1883 k	10.36
Filtered B^+	2956 k	16.42
$c\bar{c}$	9900 k	7.62
$D^0 \to K\pi$	424 k	161.3
$D^0 \to K\pi\pi\pi$	436 k	84.9
$D^0 \to K\pi\pi^0$	654 k	68.7
$D^0 \to K_S\pi\pi$	102 k	55.2

Table 4.1: Summary of Monte Carlo events that are used in this analysis.

In order to increase the amount of generic Monte Carlo events, a special filtered sample was made where all events are required to have a good lepton with momentum greater than 1.0 GeV at the Monte Carlo generator level. This requirement doesn't affect the correctly reconstructed events in our sample, since such an event must have a good lepton with a momentum greater than 1.2 GeV, but the character of background events is significantly changed, especially the fake lepton background. This filter effectively cuts the number of $D^*l\nu$ candidate events by a factor of 3 while retaining all correctly reconstructed events, a detailed comparison of what events are removed is located in appendix A. This sample is very useful for testing the results of fits vs Monte Carlo truth,

but it is not used to characterize backgrounds or derive any fit quantities related to the

measurement on data.

Dataset	Lumi (pb^{-1})	Events	Opposite Side Elec	Muon	Fake	Same Side Elec	Muon	Fake
1999-b1-s0-r8B-on0	434.7	1960	404	431	866	37	47	246
2000-b1-s0-r8B-on0	643.5	2708	616	593	1165	49	58	304
2000-b1-s1-r8C-off0	146.9	106	3	1	81	1	1	21
2000-b1-s1-r8C-on1	579.6	2505	506	497	1217	50	53	271
2000-b1-s2-r8A-on1	1654.8	7055	1515	1421	3283	143	135	827
2000-b1-s2-r8A-off1	186.0	128	6	9	89	1	1	23
2000-b1-s3-r8C-on1	284.2	1175	249	241	552	20	21	141
2000-b1-s4-r8A-off1	115.4	83	2	5	65	0	1	12
2000-b1-s4-r8A-on2	750.1	3241	711	691	1449	60	76	352
2000-b1-s5-r8D-on2	1740.8	7308	1568	1626	3283	124	152	793
2000-b1-s5-r8D-off2	330.1	227	4	13	168	0	1	48
2000-b1-s5-r8D-on3	421.2	1703	370	368	753	42	39	194
2000-b1-s6-r8A-on3	2291.3	9886	2139	2057	4510	199	186	1104
2000-b1-s6-r8A-off3	474.9	360	17	16	260	1	2	71
2000-b1-s6-r8A-on4	2379.3	10098	2230	2015	4635	221	194	1106
2000-b2-s0-r8A-on4	306.7	1331	286	277	616	18	24	154
2000-b2-s1-r8A-on4	873.4	3959	889	779	1826	87	87	448
2000-b2-s1-r8A-off4	392.3	303	8	14	219	2	2	64
2000-b2-s1-r8A-on5	740.7	3345	717	647	1622	47	61	390
2000-b2-s2-r8D-on5	895.6	3933	838	891	1771	80	69	414
2000-b2-s3-r8D-on5	1720.5	7183	1528	1465	3333	154	152	792
2000-b2-s3-r8D-off5	550.0	390	14	21	275	0	4	82
2000-b2-s3-r8D-on6	2661.4	10867	2373	2318	4967	237	195	1164
2000-b2-s3-r8D-off6	418.2	286	9	13	203	1	2	65
2000-b2-s4-r8E-on7	2214.0	8948	1877	1844	4191	179	174	1013
All On Peak	20591.8	87205	18816	18161	40039	1747	1723	9713
All Off Peak	2614.0	1883	63	92	1360	6	14	386

Table 4.2: Summary of candidates selected after applying all cuts except the vertexing selection.

4.2 Backgrounds

The most challenging aspects of this measurement are the complications due to various backgrounds for the reconstruction of a $B \to D^* l \nu$ candidate. The reconstruction is essentially a three step process. First the D^0 candidate is found in one of four decay modes. Then a soft pion candidate is added to the D^0 to create a D^* candidate and finally a lepton candidate is added to the D^* to create the $B \to D^* l \nu$ candidate. The mass difference between the D^* and D^0 is just larger than the charged π mass and forms a very narrow peak. The combinatoric background is very broad, so this peak serves as a very powerful cut against backgrounds coming from misreconstructed D^0 or D^* candidates (see figure 4.27 for a typical distribution). Unfortunately, there is no such constraint for the final leg of the reconstruction. The neutrino escapes detection and the ability to add up all the daughter energies to equal the parent is lost. This materializes in the analysis as numerous backgrounds which involve a good D^* but a bad lepton candidate in the final step of reconstructing the $B \to D^* l \nu$ candidate. There are 6 main backgrounds for this analysis:

- Combinatoric background: The D^0 is incorrectly reconstructed or the soft pion candidate + D^0 don't come from a real D^*. Cutting on D^*-D^0 mass is effective at removing backgrounds.

- Fake lepton: A real D^* is combined with a hadron instead of a lepton. This background is minimized by requiring very tight lepton identification. Lepton fake rates are well measured.

- Correlated lepton: Good D^* combined with a secondary lepton from the same B^0 or B^- parent. B^0 or $B^- \to D^* X$, $X \to l$ Y. Secondary leptons have a softer momen-

tum spectrum than primary leptons, thus a lepton momentum cut is effective.

- Uncorrelated lepton: Good D^* combined with lepton from opposite side B. B^0 or $B^- \rightarrow D^*X$, $\overline{B^0}$ or $B^+ \rightarrow Yl$. Several angular cuts are effective and help define a control sample for estimating the size of this background.

- Charged B background: A charged B can decay into a D^{**} which quickly decays into a D^* plus at least one soft pion. $B^{\pm} \rightarrow D^*l\nu + n\pi_{soft}$. This background is extremely difficult to remove. Instead, the background fraction is added a variable that will be determined during the final fitting process.

- Continuum: A $c\bar{c}$ event can have one c quark form a D^* while the other quark decays semileptonically to create a lepton. Other continuum backgrounds (uds) are completely negligible. $c\bar{c} \rightarrow D^*lX$. This background is very jetty, cuts on the D^* momentum and event topology are effective.

4.3   Reconstruction track lists

The particles that are actually observed in the BaBar detector are pions, electrons, muons, kaons, protons and gammas. All other particles are reconstructed from these basic building blocks. These particles are detected as charged tracks or neutral clusters, and they are grouped into the following lists depending on the quality and characteristics of the track/cluster [12].

- *ChargedTracks*: All reconstructed tracks from DCH or SVT. Pion mass hypothesis is used.

- *GoodTracksVeryLoose*: subset of *ChargedTracks*, must pass:

1. a distance of closest approach to the per-event beam spot of $|\Delta z| < 10$ cm, and $\sqrt{\Delta x^2 + \Delta y^2} < 1.5$ cm

2. a maximum track momentum measured in the lab frame of 10 Gev

3. a minimum number of DCH + SVT track hits ≥ 5

- *GoodTracksLoose*: subset of *GoodTracksVeryLoose*, must pass:

  1. a minimum transverse momentum of 100 MeV

  2. a minimum number of 12 track hits recorded in the DCH

- *GoodTracksTight*: subset of *GoodTracksLoose*, must pass:

  1. a distance of closest approach to the per-event beam spot of $|\Delta z| < 3$ cm, and $\sqrt{\Delta x^2 + \Delta y^2} < 1$ cm

  2. a minimum number of 20 track hits recorded in the DCH

- *GoodPhotonLoose*

  1. a minimum calorimeter energy of 30 MeV

  2. a minimum number of EMC crystals hit > 0

  3. LAT (an energy deposit shape variable) < 0.8

- *GoodNeutralLoose*

  1. a minimum calorimeter energy of 30 MeV

  2. a minimum number of EMC crystals hit > 0

  3. no LAT cut

These lists serve as a pool of candidates for creating composite particles such as the K_S or D^0.

4.4 Composition Tools

Two software packages named CompositionTools and CompositionSequences form a tool set that create all the lists of composite particles required for this analysis. CompositionTools is a set of generic particle finders which accept input lists of charged tracks, neutral tracks and/or composites plus a set of loose kinematic cut parameters to produce an output list of a particular composite particle. For example, the D^* finder requires a list of D^0 candidates plus a list of charged tracks. The output of this finder is a list of D^* candidates made from the input candidates. CompositionSequences provides a predefined set of input lists and parameters for CompositionTools that most members of the BaBar Collaboration find useful. This allows for a certain amount of uniformity and reduced redundancy between the efforts of various analysis groups.

The following composites are used for this analysis [6]:

Decay Mode	Particle Mass	Branching fraction
$\pi^0 \to \gamma\gamma$	134.977 MeV	98.798 %
$K_S \to \pi^+\pi^-$	497.672 MeV	68.61 %
$D^0 \to K\pi$	1864.5 MeV	3.83 %
$D^0 \to K\pi\pi\pi$	1864.5 MeV	7.49 %
$D^0 \to K\pi\pi^0$	1864.5 MeV	13.9 %
$D^0 \to K_S\pi\pi$	1864.5 MeV	2.7 %
$D^* \to D^0\pi$	2010.0 MeV	67.7 %
$B \to D^*l\nu$	5279.4 MeV	4.6 %

Table 4.3: Composite particles used in this analysis, provided by CompositionTools.

The π^0 candidate list is created by combining two particles from the *GoodPhotonLoose* list which satisfy the following requirements:

1. the sum of the two photon energies must be larger than 200 MeV

2. the photon pair invariant mass must be in the range of 90-170 MeV

Photon pairs which pass these cuts form the *pi0Loose* list. This list is refitted using the π^0 mass as a constraint to create the *pi0DefaultMass* list. The π^0 mass constraint improves the energy resolution, the Default list is the one used in this analysis.

The K_S candidate list, *KsLoose*, is created by combining two oppositely charged tracks from the *ChargedTracks* list and requiring that the invariant mass by four vector addition is between 300-700 MeV.

The D^0 candidates are reconstructed in four different modes and collected into a single *D0Loose* list. The kaon is selected from either the *GoodTracksLoose* or *KsLoose* list. The pion candidates are taken from the *GoodTracksVeryLoose* or *pi0DefaultMass* list. The following cuts are applied:

- $D^0 \rightarrow K\pi$: a mass window of ± 90 MeV.

- $D^0 \rightarrow K\pi\pi$: a mass window of ± 90 MeV.

- $D^0 \rightarrow K\pi\pi^0$: a mass window of ± 160 MeV.

- $D^0 \rightarrow K_S\pi\pi$: a mass window of ± 90 MeV.

The D^* candidate list, *DstarNeutralDLoose* list is created by combining a member of the *D0Loose* list with a soft pion candidate from the *GoodTracksVeryLoose* list. The *DstarNeutralDLoose* candidate must meet the following requirements:

1. a mass window of ± 500 MeV.

2. a mass window of 130-170 MeV for the D^*-D^0 mass difference.

3. a maximum momentum of 450 MeV for the soft pion.

D^* candidates where the D^0 decays into a charged kaon are called "Right Sign" if the slow pion and kaon have the expected charge correlation (K^+, π^- or K^-, π^+). Candidates not meeting this expectation are called "Wrong Sign" candidates and are discarded from the *DstarNeutralDLoose* list.

The final $D^*l\nu$ candidate is created by combining the *DstarNeutralDLoose* list with a lepton candidate from *GoodTracksTight* list. The lepton candidates are required to have a minimum momentum of 0.8 GeV in the Center of Mass frame, and "Wrong Sign" combinations of the lepton and soft pion are rejected.

4.5 Particle ID

The types of particles actually detected and used by this analysis are $K^\pm$, $e^\pm$, $\mu^\pm$, and γ. Various measured quantities are used to estimate the likelihood of a one particular particle hypothesis or another. These likelihoods can then be used to reduce the combinatoric backgrounds generated by using incorrect particle types during reconstruction.

Particle Identification (PID) uses the energy loss per distance traveled (dE/dx) in the SVT and DCH, shower shape information from the EMC plus Cherenkov light detected in the DIRC to formulate the PID hypothesis. The likelihood for each type of particle species is determined and made available in the form of a bitmap which represents a hierarchy of probabilities. PID is most important for particles which are less likely (i.e. not pions) since these particles impact the combinatoric background most heavily. This analysis makes use of PID in several different ways:

1. Kaon PID is used to reduce combinatorics while reconstructing D^0's.

2. Lepton PID is used to reduce background events while reconstructing $D^*l\nu$.

3. PID is use to help determine the flavor of opposite side B^0 to determine the mixing status of the event.

4.5.1 Kaon PID

Kaon PID is based on information from the DCH, SVT, and DIRC. The dE/dx information provided by the SVT yields better than a 2σ separation between kaons and pions up to a particle momentum of about 0.6 GeV, and up to 0.7 GeV for the DCH. For particle momentum above 1.5 GeV the DCH once again gives a better than 2σ separation due to relativistic rise. [19] (see figure 2.7 for a plot of DCH dE/dx).

The DIRC provides a measurement of the Cherenkov angle and the number of photons arriving for each charged particle passing through the DIRC's quartz bars. For particles with $momentum > mass/\sqrt{n^2 - 1}$, where n is the index of refraction for quartz, Cherenkov light will be emitted. The number of photons produced for a fixed particle path follow poissonian statistics with a central value which depends on the particle type, charge, momentum, polar angle and bar number. The particle momentum, individual photon Cherenkov angle, number of photons and photon arrival time are all combined in a simultaneous fit to calculate the Cherenkov angle of the charged track and the kaon likelihood [19].

Figure 4.2 shows the distribution of Cherenkov angle vs momentum and the kaon/pion separation achieved by the DIRC. The right side figure is based on the resolution of pions from a D^* control sample. The 2.5σ separation is a little optimistic, if both kaons and pion resolutions are measured the separation is about 2σ at 4 GeV momentum. [19]

Kaon PID is available to all analyses in the form of a bitmap with the following

levels:

- **Very Tight**: designed to keep mis-id rates below 2% up to 4 GeV.

- **Tight**: designed to keep mis-id rates below 5% up to 4 GeV.

- **Loose**: designed to keep mis-id rates below 7% up to 4 GeV.

- **Very Loose**: designed to to be highly efficient for kaons.

- **Not a Pion**: designed to maximize kaon efficiency while rejecting pions.

4.5.2 Electron PID

Electron PID makes use of information from the DCH, DIRC and EMC. The DCH once again provide dE/dx vs momentum information. The dE/dx for a electrons is peaked at $\approx$650 with a width $\sigma \approx$50. The DIRC provides information which is used only by the very tight selector. The number of measured photons should be greater than 10 and the measured Cherenkov angle is required to be within 3σ of the electron mass hypothesis. The EMC provides five measured quantities used in electron PID:

- **Lateral energy distribution (LAT)**: One of two shower shape variables:

$$LAT = \frac{\sum_{i=3}^{n} E_i r_i^2}{\sum_{i=3}^{n} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}, \ \ E_1 \geq E_2 \geq ... \geq E_n \tag{4.1}$$

where r_0 = 5 cm, the average distance between EMC crystal fronts, r_i is the distance between the ith crystal and the shower center, E_i is the energy deposited in the ith crystal and the sum is over all crystals in the shower.

LAT is essentially a normalized weighted sum of all energy deposited in EMC crystals by a shower excluding the two largest crystal deposits. Electromagnetic show-

ers deposit most of their energy in one or two crystals, so LAT is expected to be smaller for electromagnetic showers than hadronic showers [15].

- **Zernike moments (A_{nm})**: Hadronic showers tend to be more irregular than electromagnetic showers. An expansion by angular moments can exploit this fact:

$$A_{nm} = \sum_{r_i \leq R_0}^{n} \frac{E_i}{E} \cdot f_{nm}\left(\frac{r_i}{R_0}\right) \cdot e^{-im\phi_i}, \quad R_0 = 15 \; cm \tag{4.2}$$

$$f_{nm}(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)! \rho_i^{n-2s}}{s!((n+m)/2 - s)!((n-m)/2 - s)!} \tag{4.3}$$

where n,m$\geq$0, n-m even and m$\leq$n. Again E_i is the energy deposited in the ith crystal and r_i is the distance from that crystal to the shower center. A number of Zerike moments have been investigated, but the most useful one and only one currently used is the A_{42} moment [15].

- **E/p**: The measurement of energy deposited in the calorimeter divided by the associated track momentum is an excellent means of identifying electrons. When an electron enters the calorimeter, it will produce an electromagnetic shower consisting of photon and e^+, e^- pairs, which deposit the energy of the original electron in the calorimeter. An ideal calorimeter will have an E/p = 1 for an electron. Various resolution effects and shower leakage will smear this distribution in a real calorimeter. The value of E/p could even be quite a bit higher than 1 if bremsstrahlung photons enter the shower since the measured momentum will be smaller due to the bremsstrahlung, but the deposited energy will remain the same.

Muons will deposit energy in the calorimeter as a single ionizing particle and generally exit the calorimeter thus only depositing a fraction of its total energy. Hadrons sometimes interact in the calorimeter, and sometimes pass through like a muon,

but the rarely deposit all their energy like an electron [7].

- **Track-Bump separation**: The path of an associated charged track is extrapolated to the front of the calorimeter and the angular position (ϕ_{EMC}) is recorded. The angular position of the shower cluster ($\phi_{cluster}$) is also recorded. Since electromagnetic showers happen quickly, these two positions should be very close. Hadronic showers tend to develop later and thus the track-bump separation can be larger.

$$\Delta\phi = q \cdot (\phi_{EMC} - \phi_{cluster}), \quad \text{q = charge of particle} \qquad (4.4)$$

can be used to discriminate between electrons and hadrons.

- **Number of Crystals**: The total number of crystals hit during a shower should be larger 2 crystals to guard against spurious background noise.

The various levels of electron PID are summarized in table 4.4.

Category	dE/dx	$N_{crystal}$	E/p	LAT	A_{42}	$\Delta\phi$	DIRC
noCal	540...860						
veryLoose	500...1000	3	0.50...5.0	-10...10	-10...10	no	no
loose	500...1000	3	0.65...5.0	-10...10	-10...10	no	no
tight	500...1000	3	0.75...1.3	0...0.6	-10...10	no	no
veryTight	540...860	3	0.89...1.2	0...0.6	-10...0.11	yes	yes

Table 4.4: Electon PID category requirements.
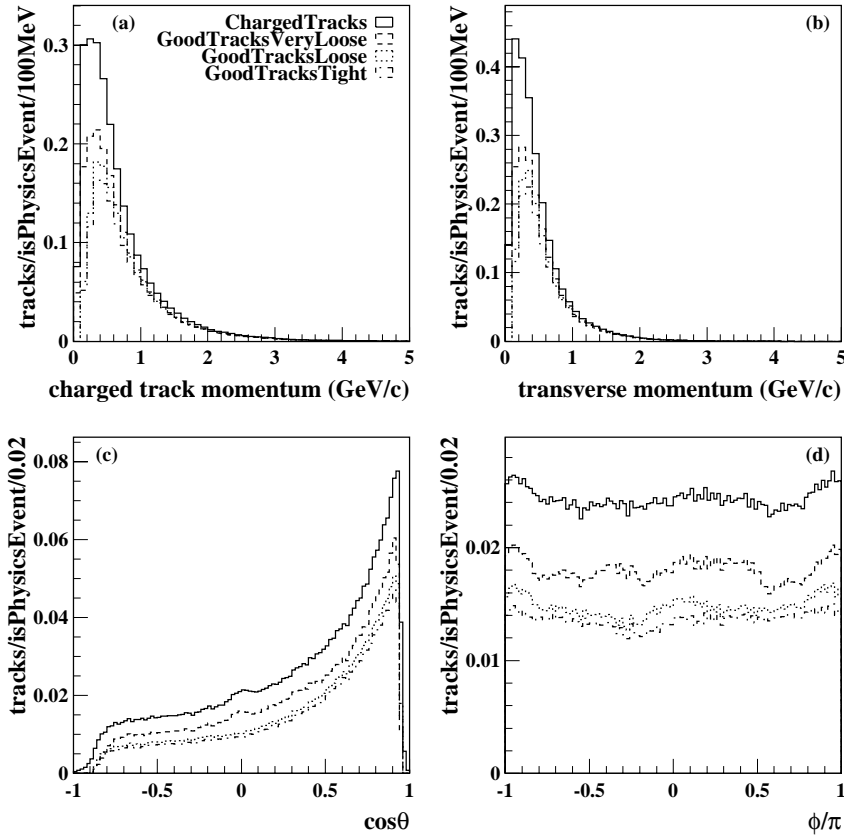
Figure 4.1: Distributions of (a) total momentum, (b) transverse momentum, (c) cos of polar angle, and (d) azimuthal angle (all measured in the lab frame) for the following lists: *ChargedTracks, GoodTracksVeryLoose, GoodTracksLoose, GoodTracksTight.* The plots are made from a typical run (number 12917), normalized to display tracks/event/bin. The tagbit used for generating these plots is called *isPhysics*, this is a very loose tagbit which essentially requires an event to have more than 2 charged tracks to remove e^+e^- and $\mu^+\mu^-$ events [12].
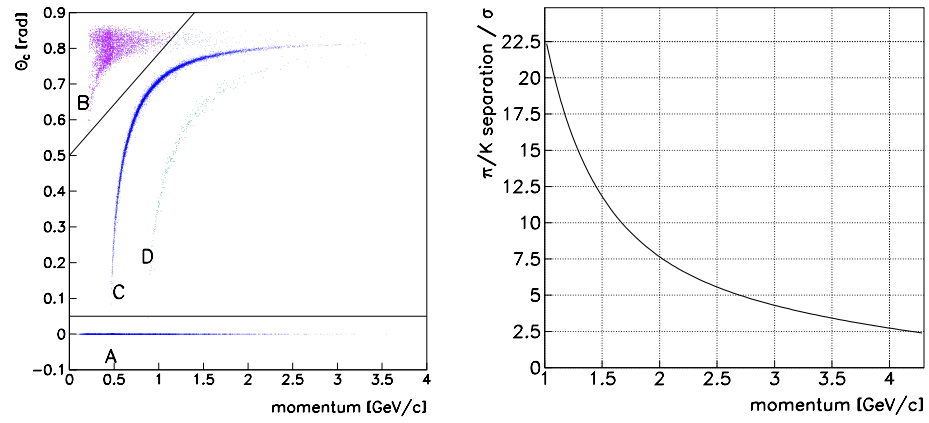
Figure 4.2: Left, the distribution of Cherenkov angle versus momentum for different particle types (the top band is pions, next one is kaon). Right, the kaon/pion separation in units of σ.

4.5.3   Muon PID

Muon PID primarily makes use of information from the IFR. The following list describes measured quantities provided by the IFR:

- N_L: The number of IFR layer hit in a cluster.

- λ: The number of interaction lengths transversed by the track. It is estimated using the track extrapolated into the IFR to the last layer hit.

- $\Delta\lambda$: The number of interaction length expected to be transversed by a muon is calculated, then the quantity $\Delta\lambda = \lambda_{expected} - \lambda$ is calculated.

- T_c: This represents the continuity of a track in a cluster. It is defined as $T_c = N_L/$(Last Layer number - First Layer number). A perfectly continuous track will have $T_c = 1$ while a track with only sporadic hits will have $T_c < 1$.

- $\overline{m}$: The average multiplicity of hit strips per layer.

- σ_m: The standard deviation of $\overline{m}$.

- χ^2_{trk}: The χ^2/d.o.f. of the IFR hit strips with respect to the track extrapolation.

- χ^2_{fit}: The χ^2/d.o.f. of the IFR hit strips with respect to a 3rd order polynomial fit of the cluster.

The amount of energy deposited in the EMC (E_{cal}) is recorded if available. Only the Minimum Ionizing category requires there to be a value for E_{cal}, all other categories drop the E_{cal} cut if there is no EMC information. Muon PID levels are summarized in table 4.5.

Category	E_{cal}	N_L	$\Delta\lambda <$	$\lambda >$	$T_c >$	$\overline{m} <$	$\sigma_m <$	$\chi^2_{trk} <$	$\chi^2_{fit} <$
Min. Ionizing	0...0.5	-	-	-	-	-	-	-	-
veryLoose	0...0.5	1	2.5	2.0	0.1	10	6	-	-
loose	0...0.5	1	2.0	2.0	0.2	10	6	7	4
tight	0.05...0.4	1	2.0	2.2	0.3	8	4	5	3
veryTight	0.05...0.4	1	2.0	2.2	0.34	8	4	5	3

Table 4.5: Muon PID category requirements.

4.6 Tag bit selection

Tag bit selection is an important process used to limit the total number of events an analysis needs to look at from the BaBar dataset. The tag bit essentially acts as a flag which indicates that a particular event has already passed a certain number of cuts. If a particular analysis uses cuts which are all as tight or tighter than a tag bit, the analysis can save a considerable amount of computer processing time by choosing to only look at events which have the tag bit set. All events which pass the L3 trigger are processed off-line to reconstruct in detail all the tracks and clusters present in each accepted event. At this time composition tools are run and various particle candidates of varying quality are created. Tag bit code uses these lists and tightens cuts made in order to reduce the number of combinatoric candidates while retaining good candidates. If at least one acceptable candidate remains in the event after this tightening process, a tag bit is set for this event. In the case of $D^*l\nu$, there are 5 tag bits of interest generated during the offline processing.

- B0ToDstarlnuLoose

- B0ToDstarlnuTight

- B0ToDstarlnuVTightElec - "very tight electron sample"

- B0ToDstarlnuVTightMuon - "very tight muon sample"

- B0ToDstarlnuVTightFake - "very tight fake lepton sample"

The very tight electron and muon tag bits represent our signal sample. The very tight fake tag bit is an important control sample with needs to be kept in order to understand lepton fake rates in our analysis. The first two tag bit are looser versions of our signal and control sample tag bits, and they are maintained in case disaster strikes and it is discovered that the last three tag bits are somehow too aggressive. The tight or loose tag bit could be used instead of the very tight tag bits if such a problem arose. The following lists show the additional constraints that each tagbit imposes.

- B0ToDstarlnuLoose

  It combines $D^*l\nu$ candidates from the DstarAllLoosePID list and lepton candidates (with no lepton identification applied) with $p_l^* > 0.8$ GeV from the GoodTracksTight list. The following D^0 decay modes are used: $K^-\pi^+$, $K^-\pi^+\pi^+\pi^-$, $K\pi^+\pi^0$ and $K_S\pi^+\pi^-$. The selection criteria are

  - π^o is fit with mass constraint.

  - Candidates with $K^-\pi^+$, $K^-\pi^+\pi^+\pi^-$ and $K\pi^+\pi^0$ modes are taken from the DstarChrgKLoosePID list with

    * D^0 from the D0ChrgKLoosePID list within ± 90 MeV mass window,

    * right and wrong sign combinations.

    The charged kaon is required to pass the SMSnotAPion selector.

  - $K_S\pi^+\pi^-$ candidates are taken from the DstarNeutralKLoosePID list with

    * D^0 from the D0NeutralKLoosePID list within ± 90 MeV mass window.

- Soft π from the GoodTracksVeryLoose list.

- $p_{\pi_{\mathrm{soft}}} < 450$ MeV.

- Raw D^* mass within 500 MeV of nominal.

- 130 MeV $< m(D^*) - m(D^0) < 170$ MeV, with mass-constraint D^0 .

- B0ToDstarlnuTight

  A refinement of the B0ToDstarlnuLoose list with additional cuts:

  - $p_l^* > 1.0$ GeV,

  - D^0 mass window narrowed to ± 40 MeV for $K^- \pi^+$, $K^- \pi^+ \pi^+ \pi^-$ and $K_S \pi^+ \pi^-$
    modes, and ± 70 MeV for $K\pi^+ \pi^0$ mode,

  - $p_{\pi_{\mathrm{soft}}}^t > 50$ MeV,

  - 0.5 GeV $< p^*(D^*) < 2.5$ GeV.

- B0ToDstarlnuVTightElec

  A refinement of the B0ToDstarlnuTight list with additional cuts:

  - $p_l^* > 1.2$ GeV,

  - D^0 mass window narrowed to ± 20 MeV for $K^- \pi^+$, $K^- \pi^+ \pi^+ \pi^-$ and $K_S \pi^+ \pi^-$
    modes, and to ± 35 MeV for $K\pi^+ \pi^0$ mode,

  - lepton candidate must pass very tight electron ID.

- B0ToDstarlnuVTightMuon

  Same as B0ToDstarlnuVTightElec except that lepton candidate must pass very tight
  muon ID.

- B0ToDstarlnuVTightFake

  Same as B0ToDstarlnuVTightElec except that lepton candidate must fail both loose electron ID and loose muon ID.

4.7 Analysis Cuts

This section will describe all the variables and cuts used in selecting $B \to D^* l \nu$ events. The presentation of the cuts will be organized to follow the logical progression of creating the D^0 composite, then the D^* and finally the $D^* l$ candidate. In many cases, a cut on a certain quality is made several times in different places (during the tagbit creation, ntuple production, ascii file production, final cut selection), the plots in this section will show the effect of the final cut on a sample.

In order to show the full range of cuts performed on the data sample, two different data sets were created:

- **No Tagbit data**: A sample roughly equivalent to 1 fb^{-1} of generic data ($B^0\overline{B}^0$ + B^+B^- + $c\overline{c}$ Monte Carlo events) was produced by Mandeep Gill which contains only the cuts imposed by Composition Tools and a lepton momentum cut of 0.8 GeV. This represents essentially the loosest cuts that can be applied. The momentum cut could be relaxed further but this adds very few signal events while introducing an extremely large number of combinatorics to the dstar-lepton list. The purpose of this sample is to study the effects of cuts applied to D^0 and D^* candidates, which can't be done when events are selected with tight tagbits.

- **Full Sample**: This is a ≈ 9 fb^{-1} sample of unfiltered $B^0\overline{B}^0$, B^+B^- and $c\overline{c}$ events (see table 4.1). All D^0 and D^* cuts plus a 1.2 MeV lepton momentum cut have been

applied to all events. This sample is useful for studying the cuts used to further refine the $D^*l\nu$ candidate list.

The great utility of Monte Carlo data is that the underlying decay is known exactly. A one to one correspondence can generally be created between detected tracks and Monte Carlo truth tracks. When a detected track has a corresponding Monte Carlo truth track, the tracks are called "MC matched". The Monte Carlo truth matching algorithm used by BaBar is extremely efficient for correctly matching charged tracks with a Monte Carlo truth partner, the same isn't true for neutral tracks.

Neutral tracks are much harder to match since there is considerably less information gathered on neutral tracks. The primary detector for neutrals is the EMC while charged tracks can leave information in the SVT, DCH, DIRC, EMC and even IFR for muons. This weakness fortunately has very little effect on this analysis. The only possible problem occurs in the $D^0 \rightarrow K\pi\pi^0$ mode, but there exist enough information in the parent D^0 and charged daughters to overcome any possible problems in MC matching. The neutrino is never MC matched since it is never detected.

4.7.1 D^0 **Cuts**

The first composite particle created is the D^0. The first cut made on the D^0 candidates is a mass cut. Figure 4.5 shows the D^0 masses for all candidates in red, the blue plot indicates D^0's which are matched to a real Monte Carlo D^0, and the black dotted lines show the value of the final D^0 mass cut value, a list of all cut values is complied in the final cut section.

The quality of the vertex fit is also used to remove background events. The χ^2 and degrees of freedom of the fit are converted into a probability and events with a

probability less then 0.001 are rejected, see figure 4.6. Tight Kaon PID is used to eliminate combinatoric D^0's in the $D^0 \rightarrow K\pi\pi$ and $D^0 \rightarrow K\pi\pi^0$ modes. NotAPion Kaon PID is used in the $D^0 \rightarrow K\pi$ mode, since it is already very clean. Kaon PID is not available for the K_S mode. Figure 4.7 shows the number of events eliminated by the Kaon PID cut while figure 4.8 shows the number of MC matched events lost due to this cut.

The decay modes with a π^0 or K_S have additional cuts to reduce extra backgrounds produced by these neutral particles. The $D^0 \rightarrow K\pi\pi^0$ decay has several resonance states which contribute to the decay. The most important resonance is $D^0 \rightarrow K^-\rho^+$, but the $D^0 \rightarrow K^{*-}\pi^+$ and $D^0 \rightarrow K^{*0}\pi^0$ are also significant. The Dalitz plot densities are calculated with the three-body decay kinematic variables, $m^2_{K^-\pi^+}$, $m^2_{K^-\pi^0}$, $m^2_{K^-\pi^+\pi^0}$. The Dalitz plot distribution at generator level and contours is shown in figure 4.3. This distribution is then convoluted with $m^2_{K^-\pi^+}$ and $m^2_{K^-\pi^0}$ resolution functions to produce the final density plot. The plot is normalized such that the largest value possible is 1. Events with a value outside the 0.1 contour are rejected. The same procedure is applied to the $D^0 \rightarrow K_S\pi\pi$ mode. The result of this cut is shown in figure 4.10 and figure 4.12.

Both the π^0 and K_S also have a vertex on the daughters of each neutral particle, shown in figure 4.9 and figure 4.11. Figure 4.4 shows the mass distributions for the π^0 and K_S.

These cuts significantly reduce the number of combinatoric D^0's while retaining the majority of good candidates.
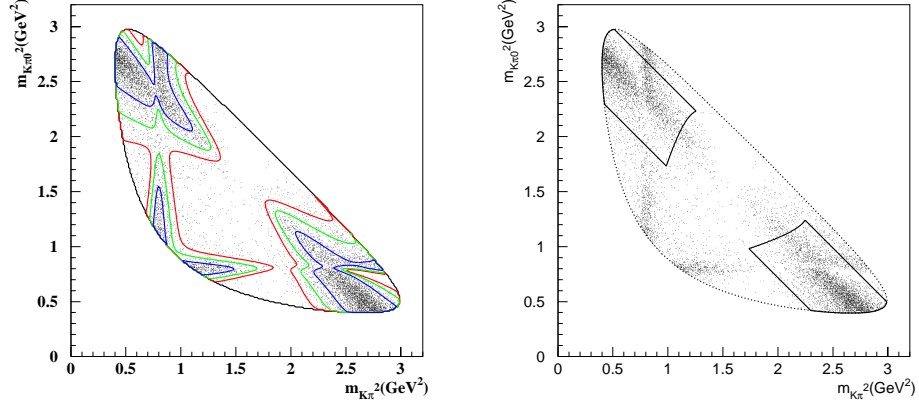
Figure 4.3: Left: $D^0 \to K\pi\pi^0$ Dalitz distribution at the generator level without convolving the distribution with resolution functions. Right: The region which falls inside the typical cut based region of $|m(\pi^+\pi^-) - m(\rho)| <$250 MeV and $|cos\theta^*_{K\pi^+}| >$0.4 in the $\pi^+\pi^-$ rest frame.

D^0 cut	$K\pi$	$K\pi\pi\pi$	$K\pi\pi^0$	$K_S\pi\pi$
Mass cut(1864.5 MeV)	±17 Mev	±17 MeV	±34 MeV	±17 MeV
D^0 vertex	0.001	0.001	0.001	0.001
Kaon PID	Not A Pion	Tight	Tight	none
Dalitz cut			0.1	0.1
K_S, π^0 vertex			0.01	0.01
K_S, π^0 mass cut(MeV)			134.977$\pm$15.75	497.672$\pm$15

Table 4.6: Summary of D^0 cuts.
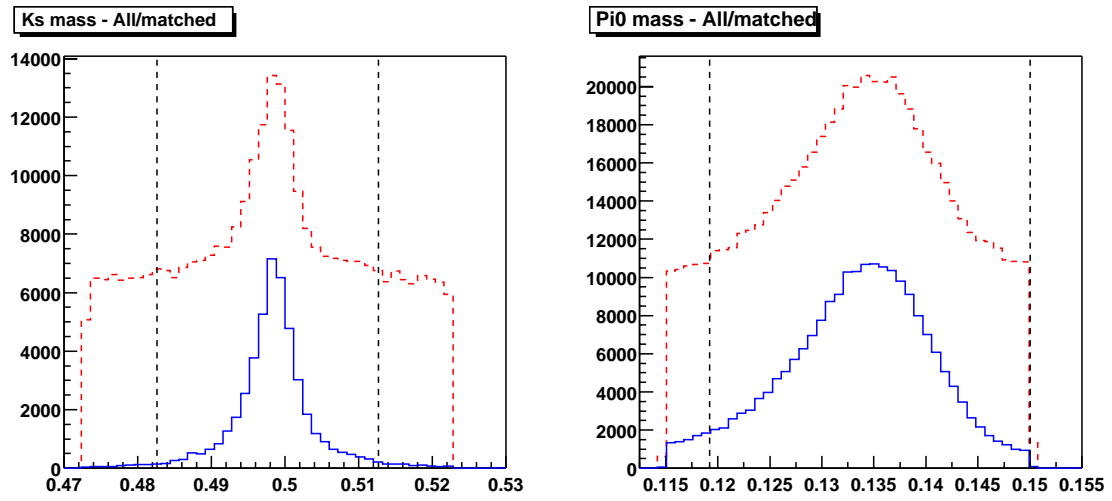
Figure 4.4: Left: K_S mass plot, the dash red curve is all events, solid blue curve is MC matched events. Right: π^0 mass plot, the dash red curve is all events, the solid blue curve is MC matched events.
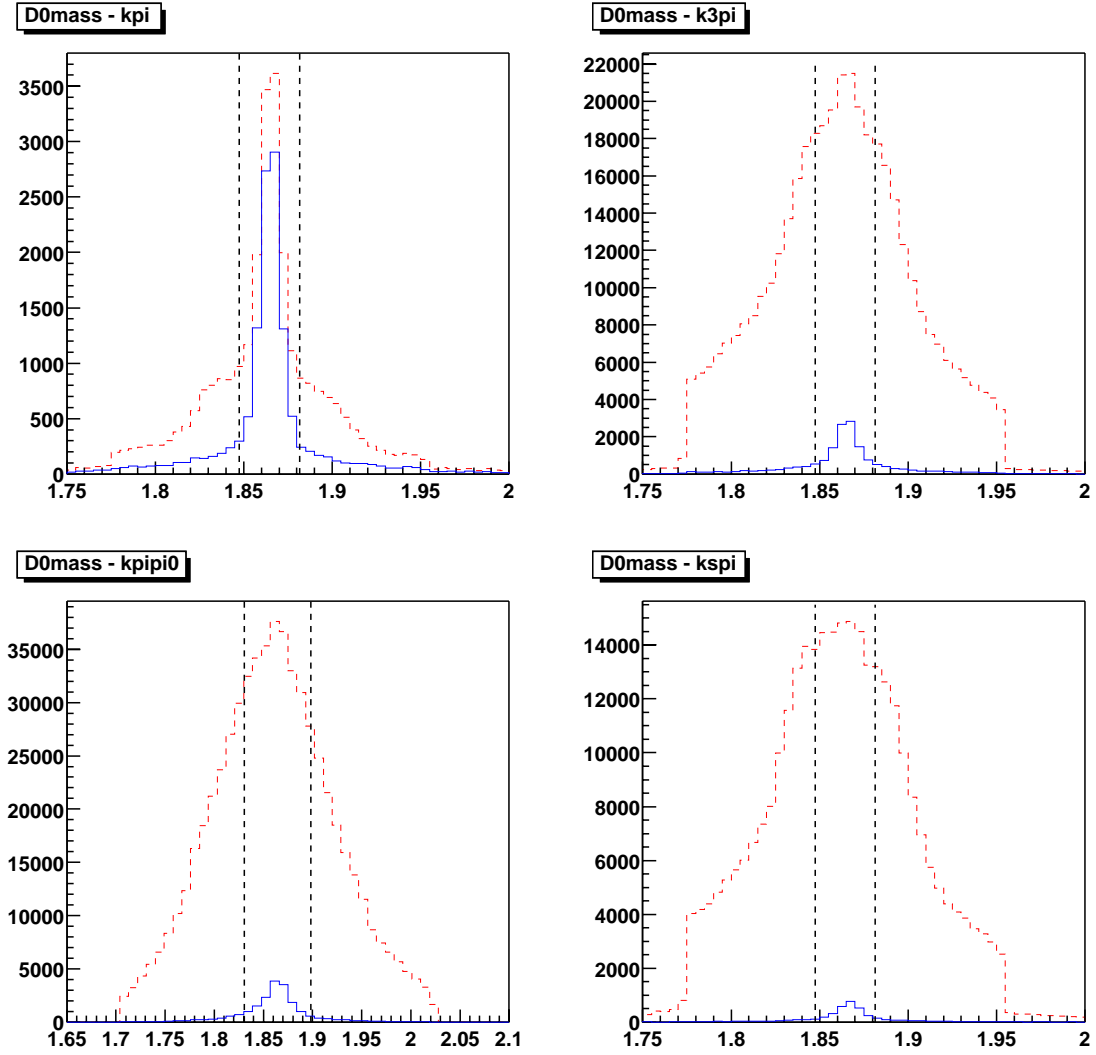
Figure 4.5: A plot of D^0 masses for all D^0 candidates (red dash) and MC matched candidates (solid blue). The final cut values are shown by the black vertical lines.
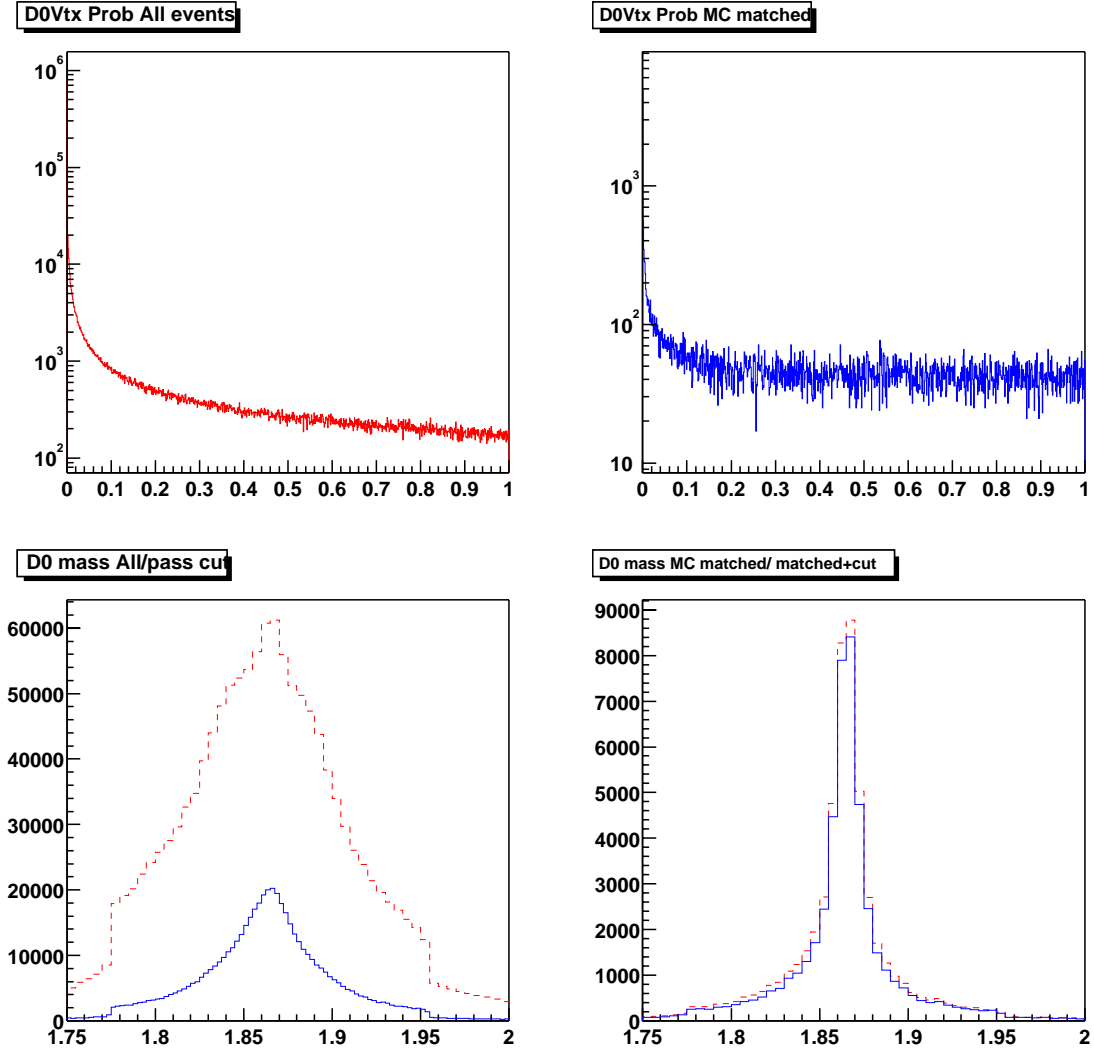
Figure 4.6: Top: Plots of D^0 vertex probabilities for all D^0 candidates and MC matched candidates. Bottom,left: This plot shows the effect of the D0Vtx cut on all events - the red dash curve is all events while the solid blue curve is events passing the D^0 vertex cut. Bottom,right: This shows the effect of the cut on MC matched events. The red dash curve is all MC match event before the cut, the solid blue curve is MC matched events after the cut.
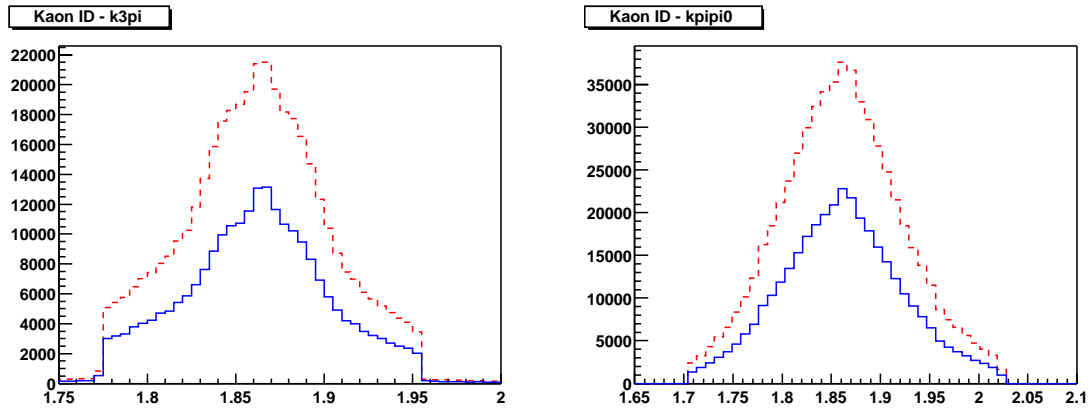
Figure 4.7: The effect of applying Kaon id to the the D^0 samples. The red dash curve shows all events, the solid blue curve is all events which pass Kaon id. There is no Kaon id for K_S, the $k\pi$ mode uses the "not a pion" requirement which has virtually no effect at all. The $K^-\pi^+\pi^+\pi^-$ and $K\pi^+\pi^0$ modes use the "tight kaon" PID, these two modes are shown above.
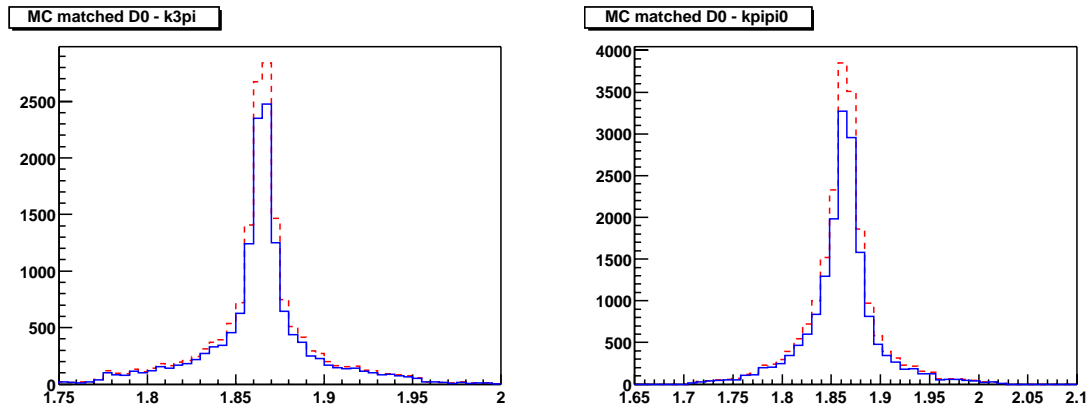
Figure 4.8: Left: The effect of applying Kaon id to the the D^0 samples. The red dash curve shows all MC matched events, the solid blue curve shows the MC matched events remaining after the Kaon id cut.
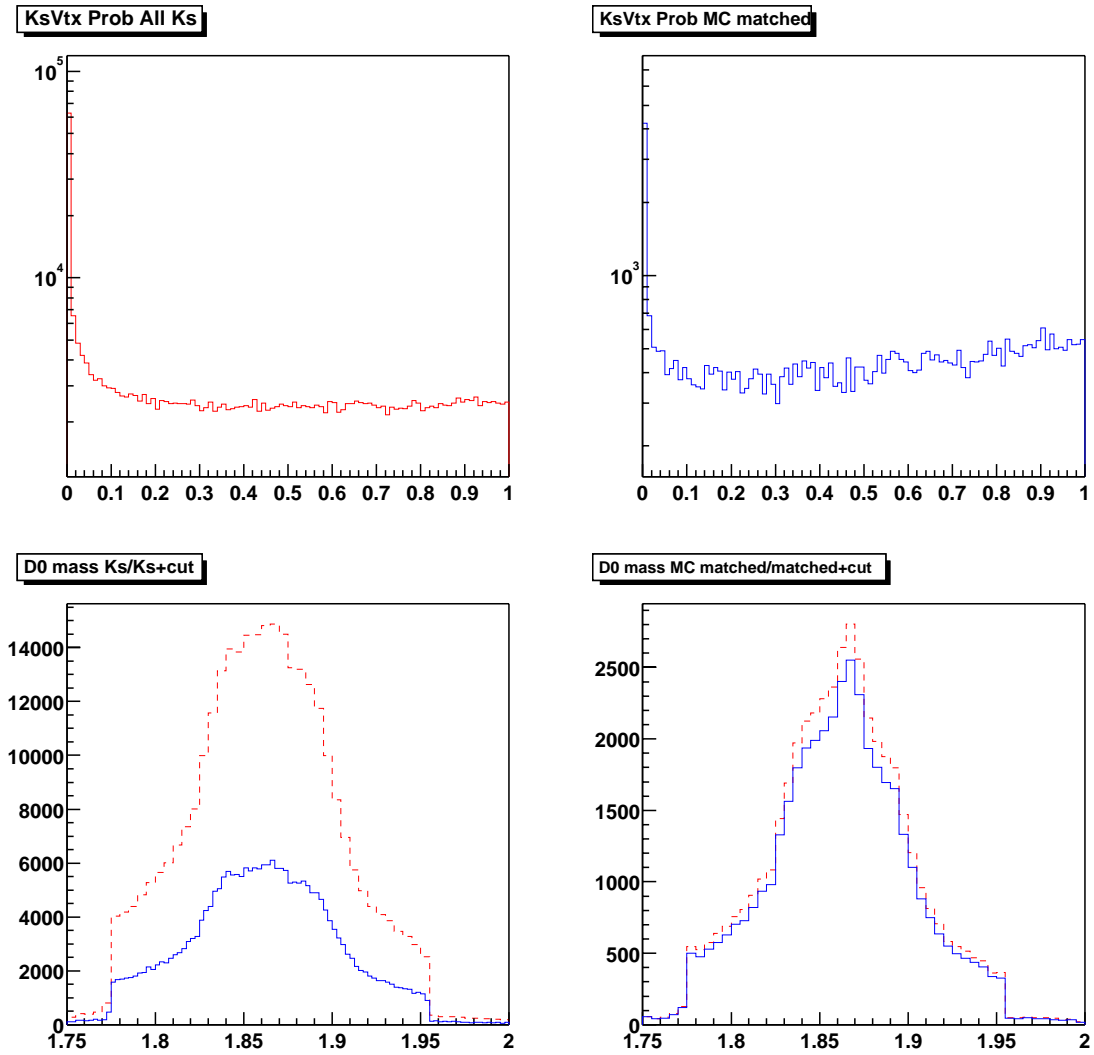
Figure 4.9: Top: The left plot shows all K_S events, the right plot shows all MC matched K_S events. The vertex cut applied on K_S is prob(χ^2)>0.01 (the first bin on both plots). Bottom left: shows all events (red dash) prior to the vertex cut and events passing the vertex cut (solid blue). Bottom right: the red dash curve is all MC matched K_S events prior to the cut, the solid blue curve is all MC matched events which pass the vertex cut.
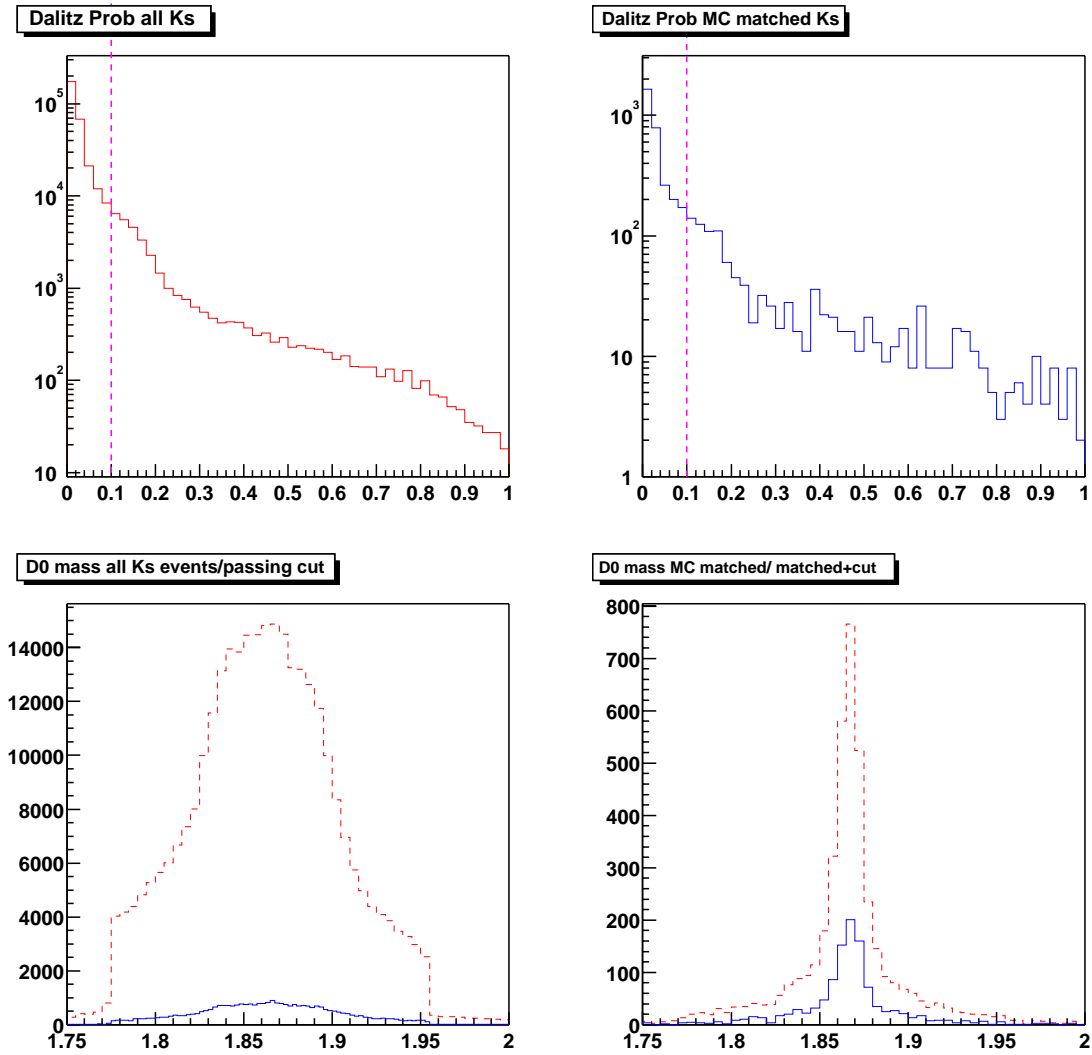
Figure 4.10: Top left: Dalitz values for all K_S events. Top right: Dalitz values for all MC matched K_S events. Bottom left: the red dash curve is all K_S events, the solid blue curve is all K_S events passing Dalitz$>$0.1. Bottom right: the red dash curve is all MC matched K_S events, the solid blue curve is all MC matched K_S events which pass Dalitz cut. The cut is $>$0.1
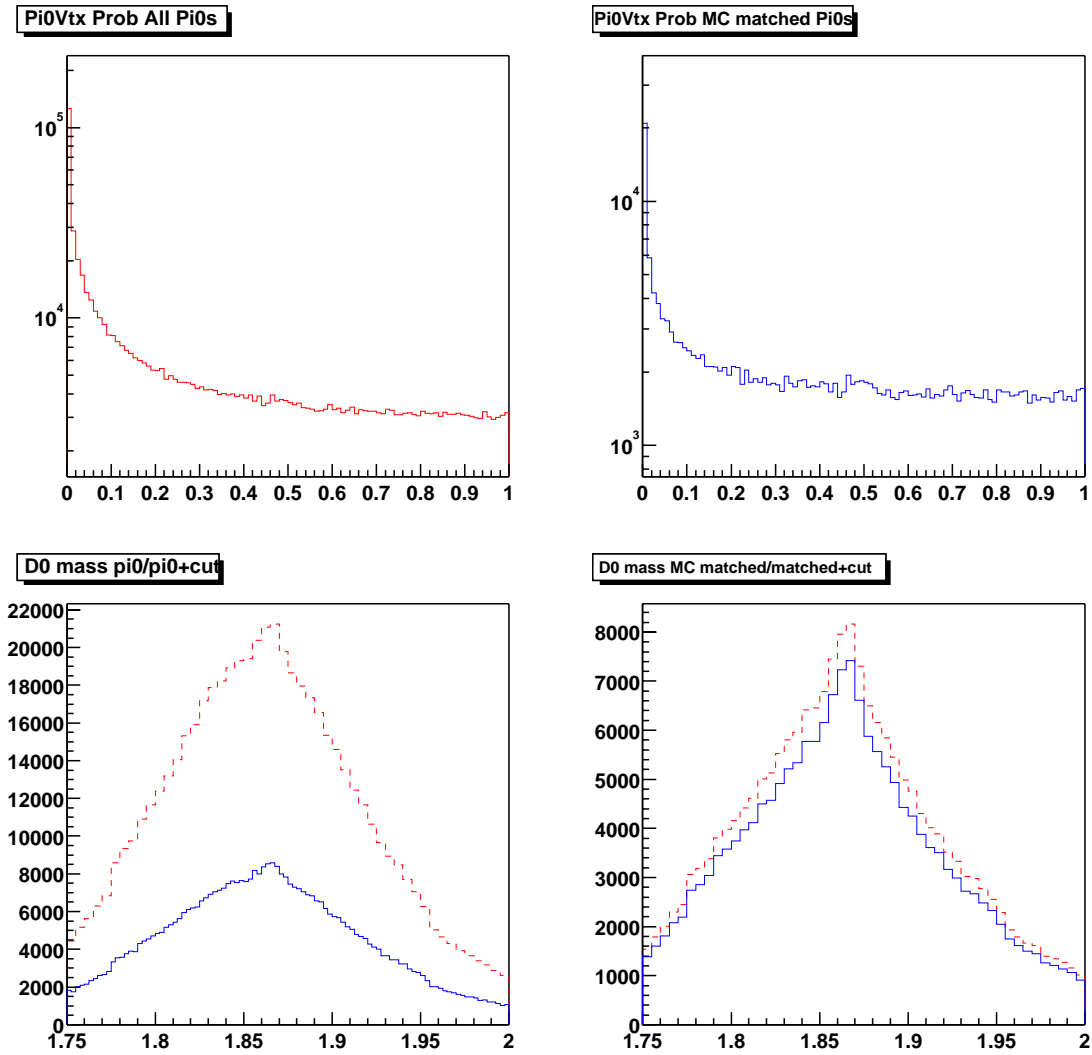
Figure 4.11: Top: The left plot shows all π^0 events, the right plot shows all MC matched π^0 events. The vertex cut applied on π^0 is prob(χ^2)>0.01 (the first bin on both plots). Bottom left: shows all events (red dash) and events passing the vertex cut (solid blue). Bottom right: the red dash curve is all MC matched π^0 events, the solid blue curve is all MC matched events which pass the vertex cut.
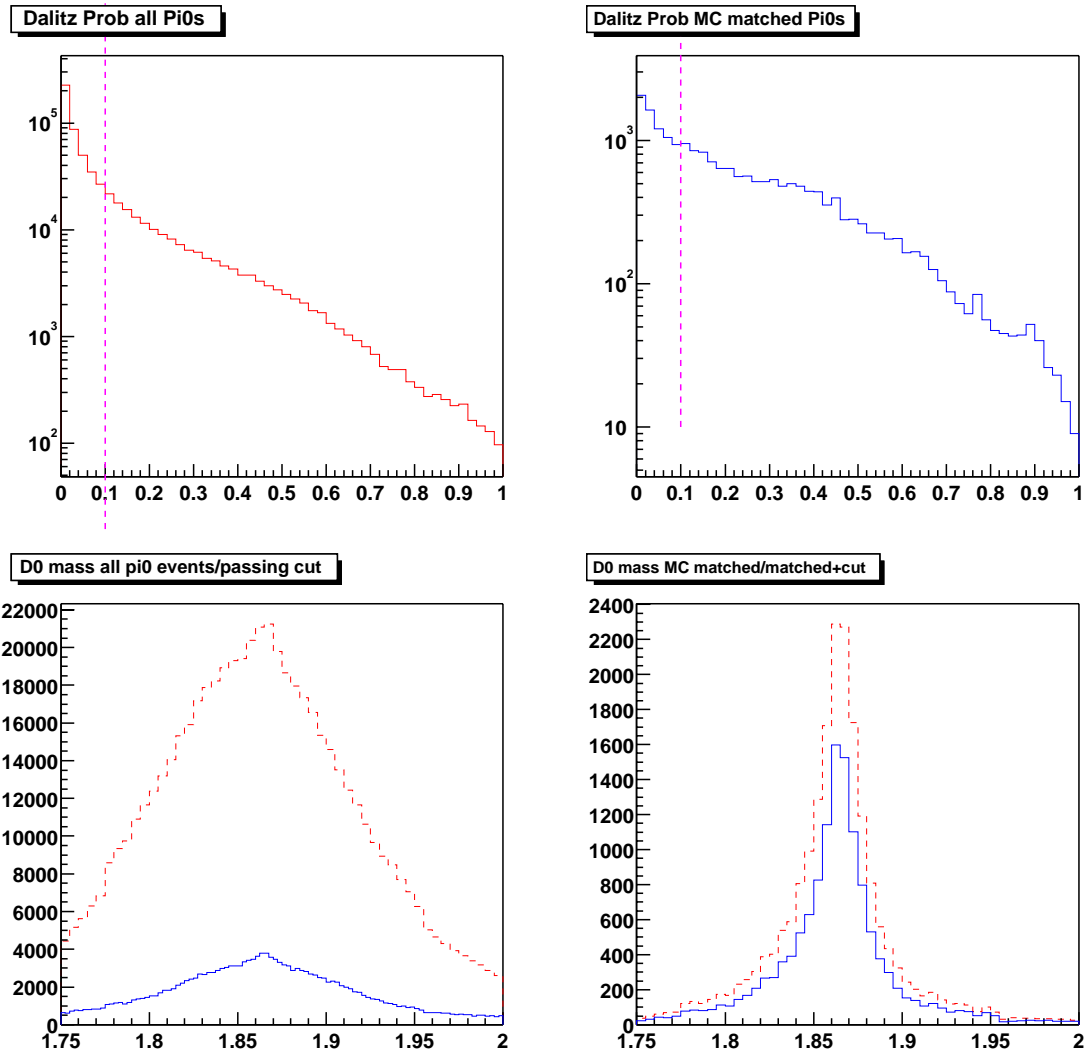
Figure 4.12: Top left: Dalitz values for all π^0 events. Top right: Dalitz values for all MC matched π^0 events. Bottom left: the red dash curve is all π^0 events prior the Dalitz cut, the solid blue curve is all π^0 events passing Dalitz>0.1. Bottom right: the red dash curve is all MC matched π^0 events prior to the Dalitz cut, the solid blue curve is all MC matched π^0 events which pass Dalitz cut. The cut is >0.1

4.7.2 D^* **Cuts**

This stage of reconstruction involves adding a single charged particle from the *GoodTracksVeryLoose* list to the D^0 candidate. Since the D^0 mass plus π mass almost add up to the D^* mass (D^* mass - D^0 mass - π mass ≈ 5 MeV), the D^0 and π are produced nearly at rest in the D^* frame. The D^0 and π have almost the exact same velocity as the parent D^* which means that the D^0 and π split the parent momentum in the same ratio as the ratio of their masses ($1865/139\approx13$). Thus the π will have a small momentum in the lab frame, which is where the label "soft pion" comes from.

There are two quality cuts made to improve the purity of the D^* sample. The first is a cut of 0.05 GeV on the minimum transverse momentum of the soft pion, as shown in figure 4.13. The second is a cut on the D^* momentum in the center of mass frame, shown in figure 4.14. The lower cut of 0.5 GeV is close to the kinematic lower limit for good D^*'s, the 2.5 GeV upper limit cuts out events which have a good D^* but not coming from $B \to D^*l\nu$.
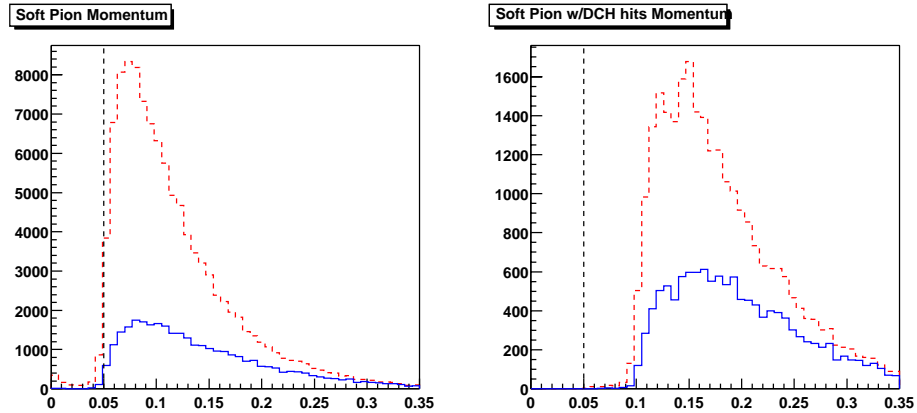
Figure 4.13: Left: the momentum distribution of the soft pion. The red dash curve represents all D^* candidate, the solid blue curve is MC matched D^*'s. Right: Same as right side except the soft pion candidate is required to have at least 5 DCH hits.
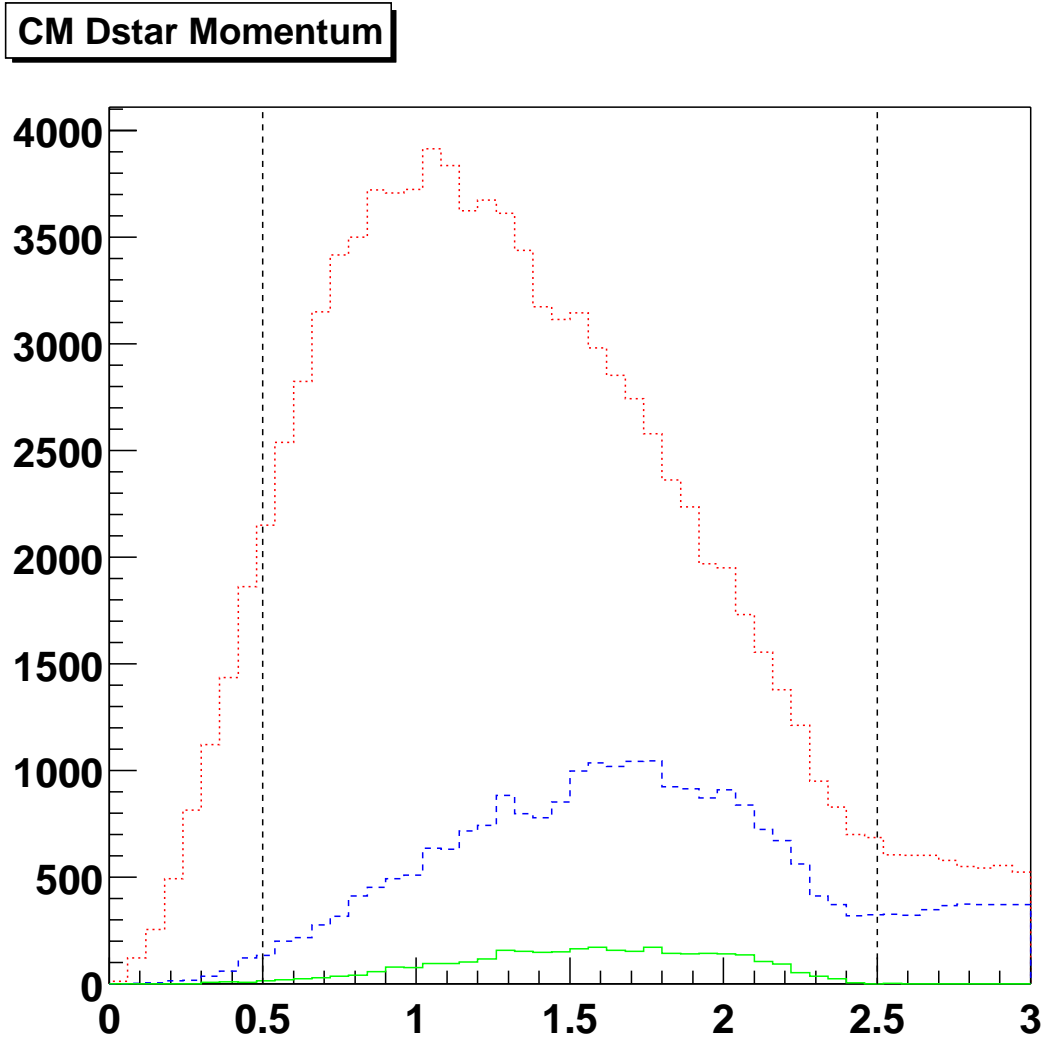
Figure 4.14: The momentum of the D^* in the center of mass reference frame. The red dotted curve represents all D^* candidates, the blue dash curve is all MC matched D^* candidates while the solid green curve shows all MC matched $D^*l\nu$ candidates (this shows the effectiveness of the upper limit cut).

4.7.3 $D^*l\nu$ Cuts

The first four plots of this section use the special 1 fb^{-1} sample with all D^0 and D^* cuts applied, the remaining plots show variables which are not affected by Dstarlnu-User ntuple level cuts, which means that the full 9 fb^{-1} sample can be used to show the effect of the remaining cuts.

There are several new variables which help improve the quality of the signal sample:

- **Thrust**: The angle, θ_{thrust}, is the angle between the thrust axis of the D^*l candidate and the thrust axis of the rest of the event. Continuum background events tend to be very jetty and have a θ_{thrust} close to one.

- **nBadTracks**: All final charged tracks in the event used to construct the $D^*l\nu$ are required to pass $|\Delta z| < 3$ cm, and $\sqrt{\Delta x^2 + \Delta y^2} < 1$ cm (the impact parameter cut of *GoodTracksTight* but no requirement on DCH hits). The number of charged tracks used to create the $D^*l\nu$ candidate not passing these requirements is recorded.

- **cosθ_{B,D^*l}**: The cos of the angle between the B direction and the D^* + lepton. This angle comes from using the 4-vector equation $P_B = P_{D^*l} + P_\nu$. If the P_{D^*l} term is subtracted from both sides then both sides squared yields:

$$0 = m_B^2 + m_{D^*l}^2 - 2(E_B E_{D^*l} - |p_B||p_{D^*l}|cos\theta_{B,D^*l}) \qquad (4.5)$$

$$cos\theta_{B,D^*l} = \frac{2E_B E_{D^*l} - m_B^2 - m_{D^*l}^2}{2|p_B||p_{D^*l}|} \qquad (4.6)$$

Since each term is either known ahead of time (m_B, E_B, p_B) or measured (m_{D^*l}, p_{D^*l}) this angle should only take on physical values ($|cos\theta_{B,D^*l}| \leq 1$). Due to resolution effects the actual cut is $|cos\theta_{B,D^*l}| \leq 1.1$.

- **cosθ_{B,D^*-l}:** The same calculation as cosθ_{B,D^*l} except the three momentum of the lepton is "flipped", i.e. $\vec{p}_{lepton} \rightarrow -\vec{p}_{lepton}$. This should not in principle affect the distribution of uncorrelated lepton events while other types of events are significantly reduced. This variable is used to obtain an enriched sample of flip lepton events for background characterization.

- **cosθ_{D^*l}:** This is the angle between the D^* and lepton candidate. Because the B^0 is a spin zero particle the total spin of the D^*,l,ν system must be zero. With the D^* being spin 1, the lepton-neutrino system will try to align its spin angular momentum to cancel the total angular momentum projected on the z axis. With the helicity of a neutrino being fixed (left-handed neutrinos, right handed anti-neutrinos) the lepton tends to come out back to back from the D^* with a harder spectrum than the neutrino. A number of backgrounds will unfortunately have the same topology; ccbar events will have a D^* and lepton essentially back to back as will fake lepton events (B^0 two body decay). The one type of background this cut is effective against is the uncorrelated lepton, which has a flat distribution of cosθ_{D^*l} before the cosθ_{B,D^*l} cut is made. The cosθ_{D^*l} is required to be less than 0 for the signal sample.

Good $D^*l\nu$ candidates are required to pass the cuts shown in table 4.7.

Figure 4.15 shows the vertex cut applied to the $D^*l\nu$ vertex and the effect on combinatoric backgrounds. Figure 4.16 shows the center of mass momentum of the lepton candidate for all $D^*l\nu$ candidates and the MC matched $D^*l\nu$ candidates. This cut is very effective at removing large combinatorics and removing cascade or secondary leptons which are real leptons with a softer spectrum than primary leptons.

Figure 4.17 shows the effect of requiring all tracks to pass the good tracks cri-

D^* cut	Cut value
Slow pion momentum	0.05< p < 0.45 GeV
Momentum D^*	$0.5 < p^*_{D^*} < 2.5$ GeV
$D^*l\nu$ cut	
$D^*l\nu$ vertex probability	0.01
P^*_{lepton}	1.2 GeV
Charged Tracks	$|\Delta z| < 3$ cm, $\sqrt{\Delta x^2 + \Delta y^2} < 1$ cm
Lepton PID	Very Tight electron or muon PID
Thrust cut	$|\cos\theta_{thrust}| \le 0.85$
$\cos\theta_{D^*l}$	$\cos\theta_{D^*l} < 0$
$\cos\theta_{B,D^*l}$	$-1.1 \le \cos\theta_{B,D^*l} \le 1.1$

Table 4.7: Summary of D^* and $D^*l\nu$ cuts.

teria. There is actually very little noticeable effect which indicates that the previous cuts and requirements have eliminated most tracks that don't appear to originate from the beam spot. Figure 4.18 shows the impact of using very tight lepton PID. This cut remove large amounts of background while retaining most of the MC matched candidates.

Figure 4.19 is made from the ≈ 9 fb^{-1} generic+ccbar sample with all Dstarlnu-User cuts (everything up to this point) applied. The plots also contain a cut on the D^* - D^0 mass of 0.1454$\pm$0.0025 GeV, which is considered to the signal region in D^* - D^0 mass space. The rest of the plots in this section (figures 4.20 through 4.25) are made with the DstarlnuUser cuts, the D^* - D^0 mass window cut and a cut of $|\cos\theta_{thrust}|$ <0.85. Since the two angular cuts are correlated, cuts are shown individually and on a 2-D scatter plot.
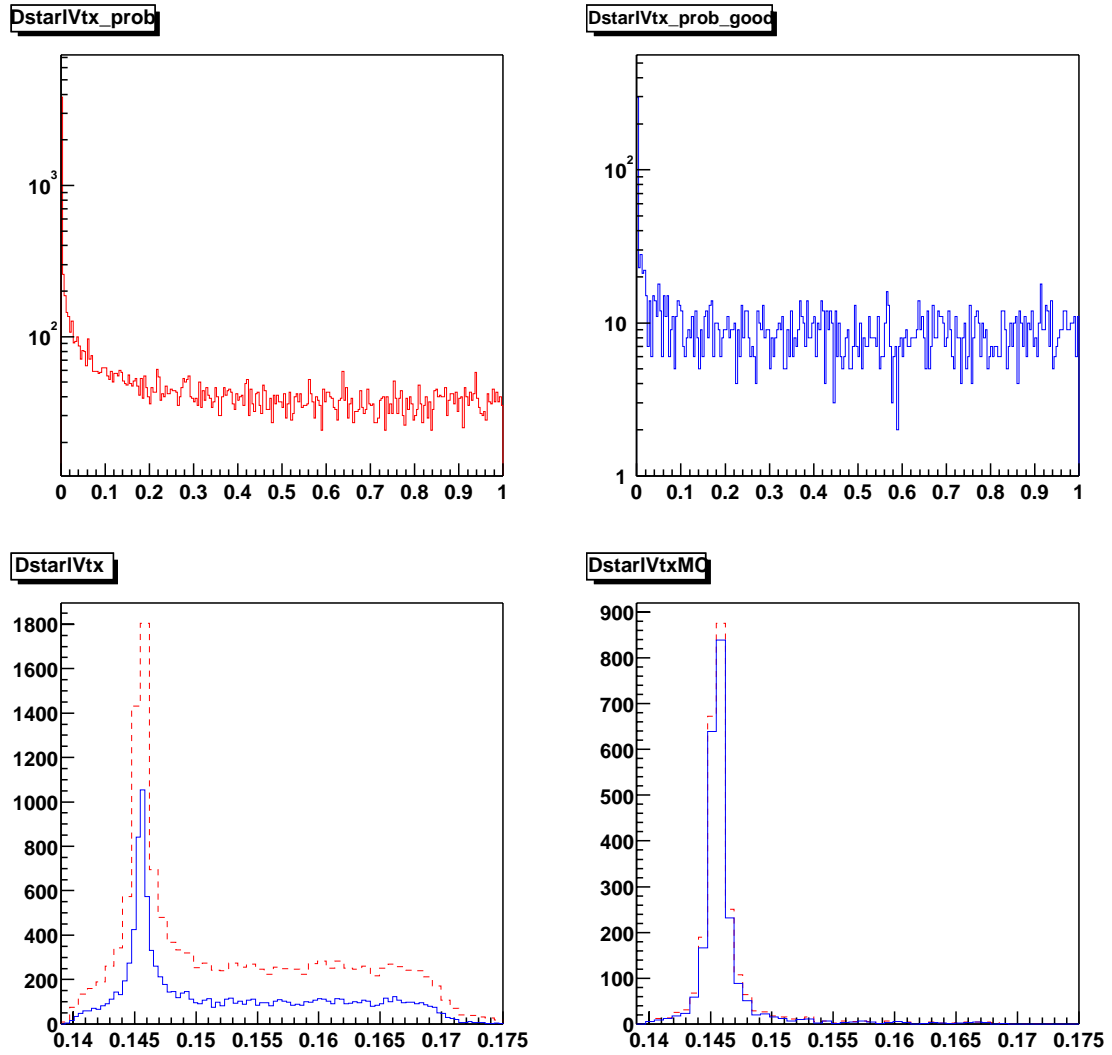
Figure 4.15: The vertex quality cut applied to the D^*l vertex. Top: The left plot represents all events while the right plot corresponds to MC matched events. Bottom: the left shows all events (red dash) and events passing the cut (solid blue) while the right shows MC matched events (red dash) and MC matched events passing the vertex cut (solid blue).
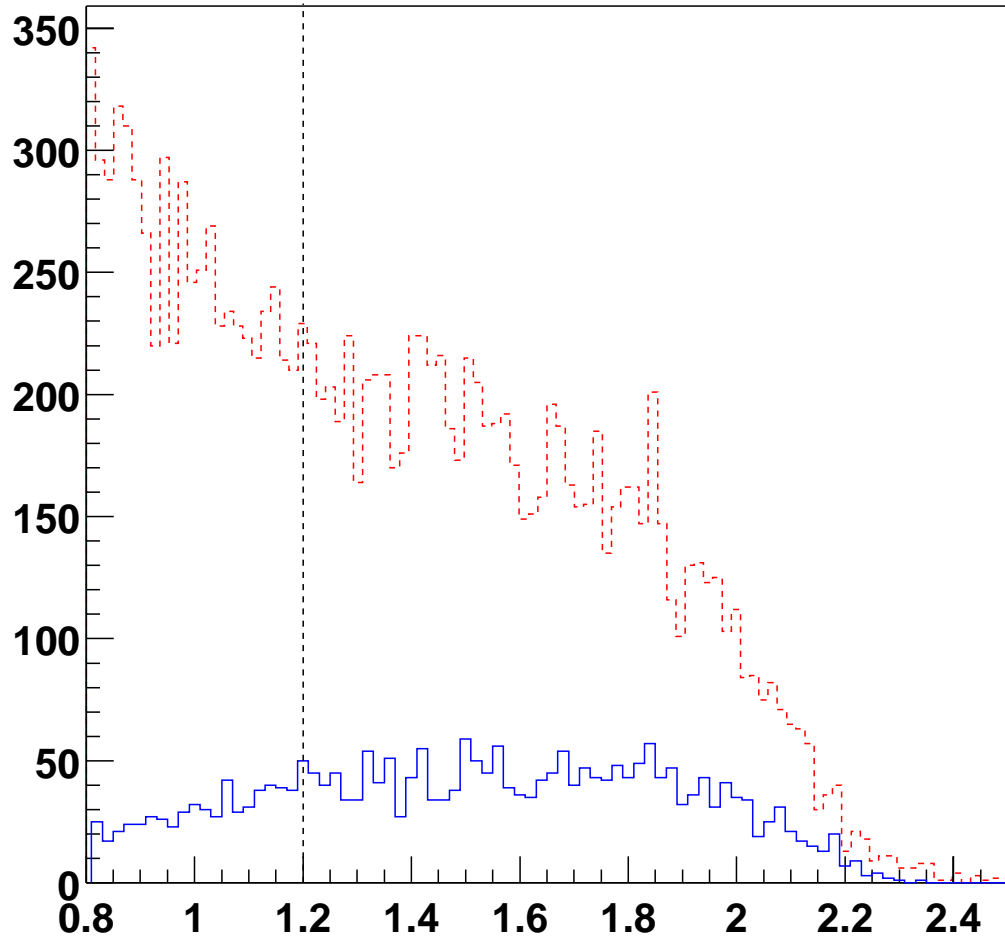
Figure 4.16: The momentum of the lepton candidate in the center of mass frame. The red dash curve shows all candidates while the solid blue curve shows MC matched $D^*l\nu$ candidates.
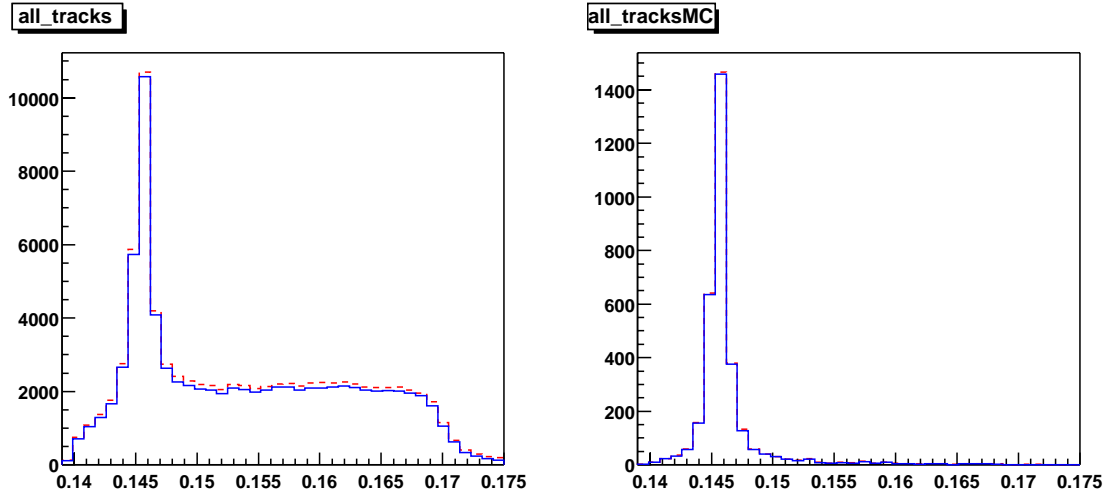
Figure 4.17: This plot shows the effect of requiring all tracks to pass the requirement of "no bad tracks". The left is all events (red dash) and all events passing the cut (solid blue), the right is all MC matched tracks (red dash) and MC matched tracks passing the cuts (solid blue).
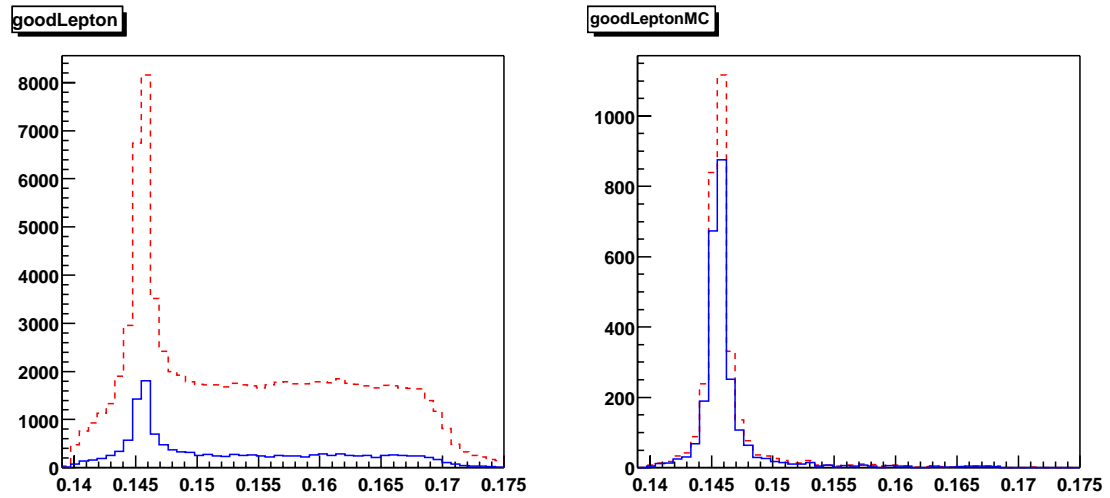
Figure 4.18: Left: the red dash curve shows all events while the solid blue curve shows all events passing the good lepton cut. Right: the red dash curve shows all MC matched events and the solid blue curve shows all MC matched events passing the good lepton cut.
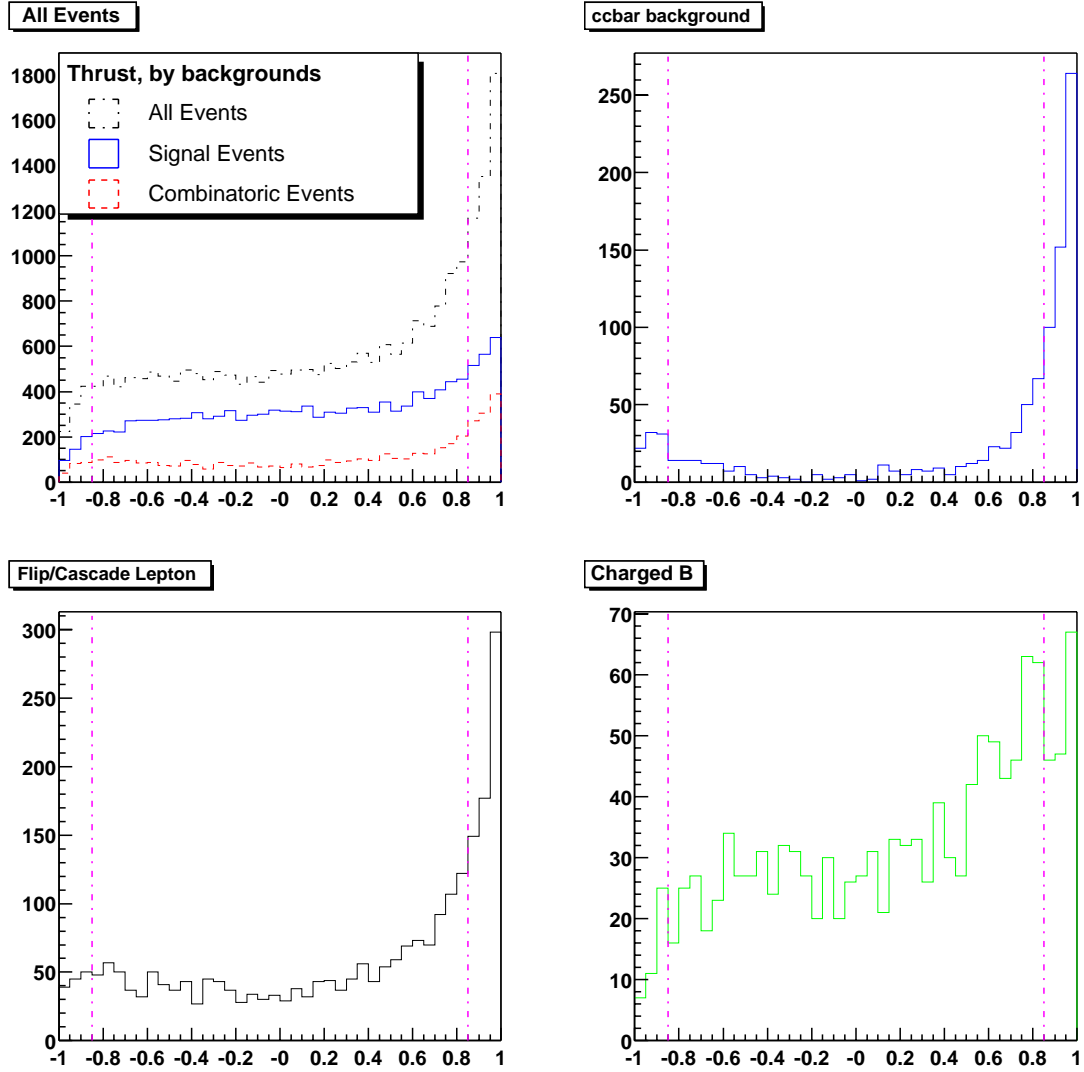
Figure 4.19: The distribution of the thrust variable broken down by different types of backgrounds. Upper left: the black line shows the thrust distribution after all D^0 and D^* cuts have been made (there is a D^* - D^0 mass cut applied). The blue line shows MC matched $D^*l\nu$ events, the red show the combinatoric background (it is modest because of the mass cut). The other plots show the ccbar, flip/cascade lepton and charged B plots.
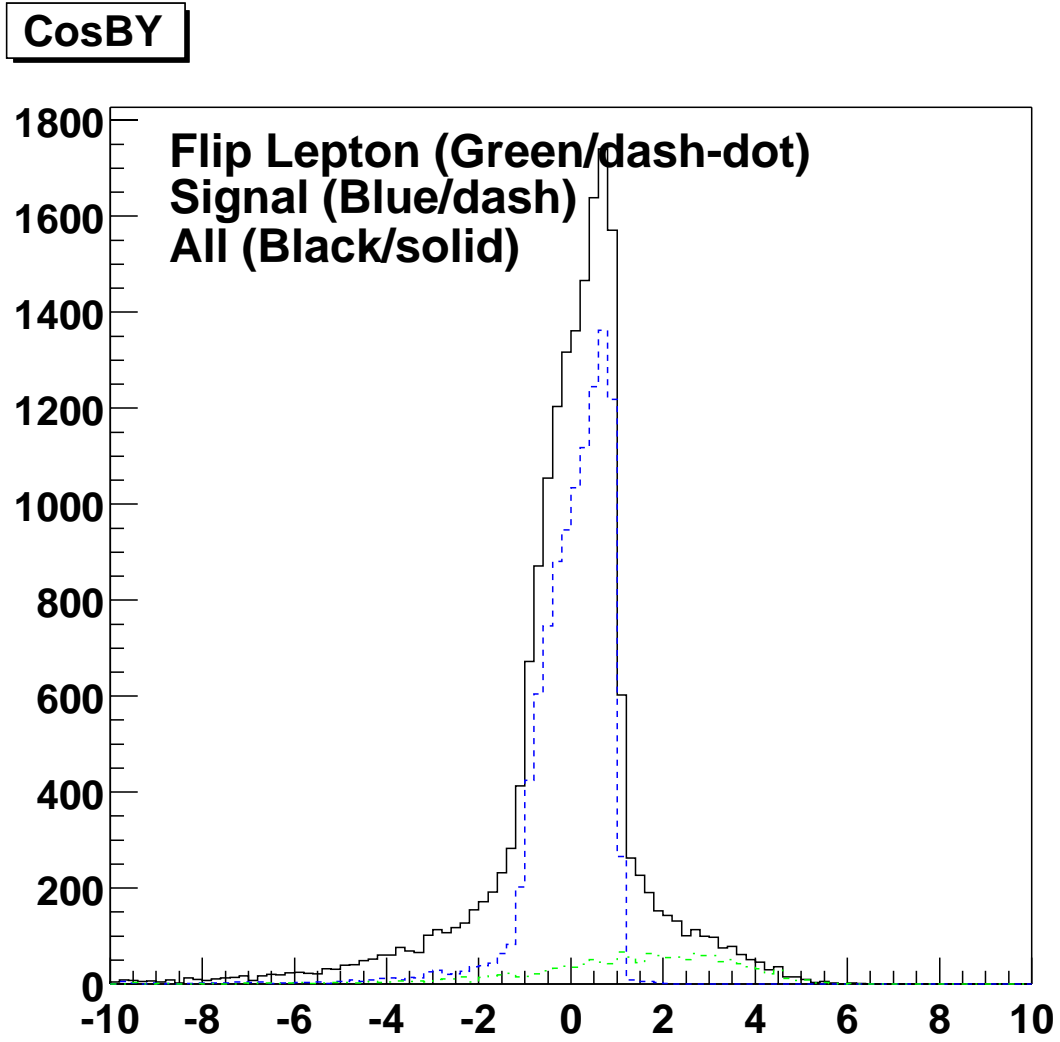
Figure 4.20: A plot of $\cos\theta_{B,D^*l}$ for all events, MC matched $D^*l\nu$ signal events and flip lepton events (i.e. a real D^* from one B and a real lepton from the other).
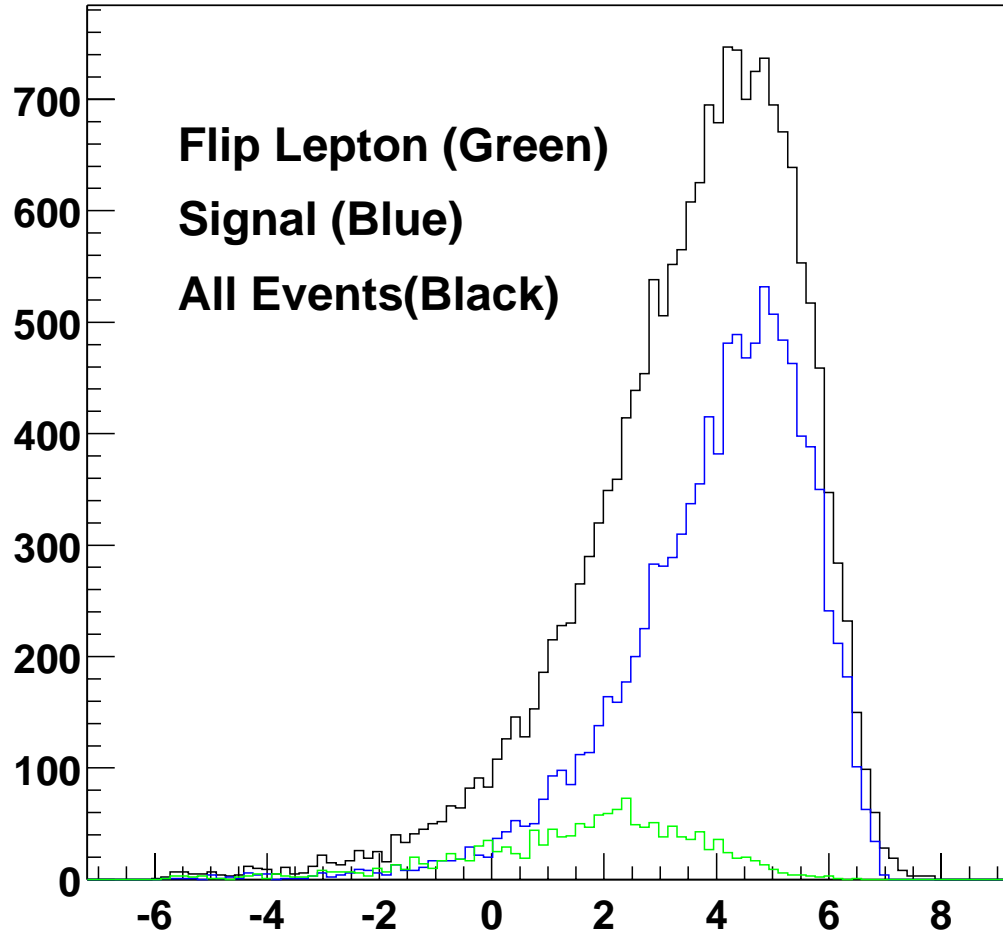
Figure 4.21: A plot of $\cos\theta_{B,D^*-l}$. Much of the signal and background events shift out of the cut region (± 1.1) except flip lepton events.
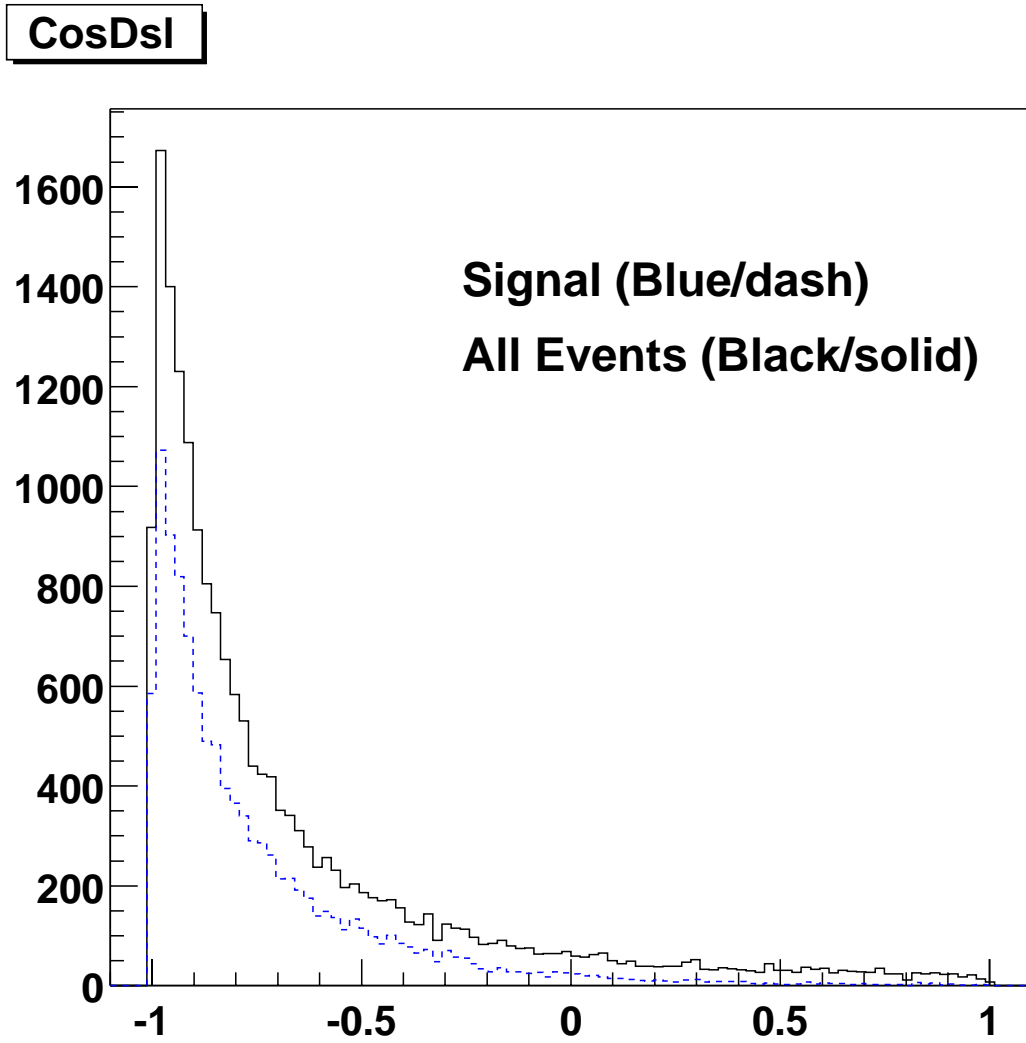
Figure 4.22: A plot of $\cos\theta_{D*l}$, the signal region is defined as $\cos\theta_{D*l} < 0$.
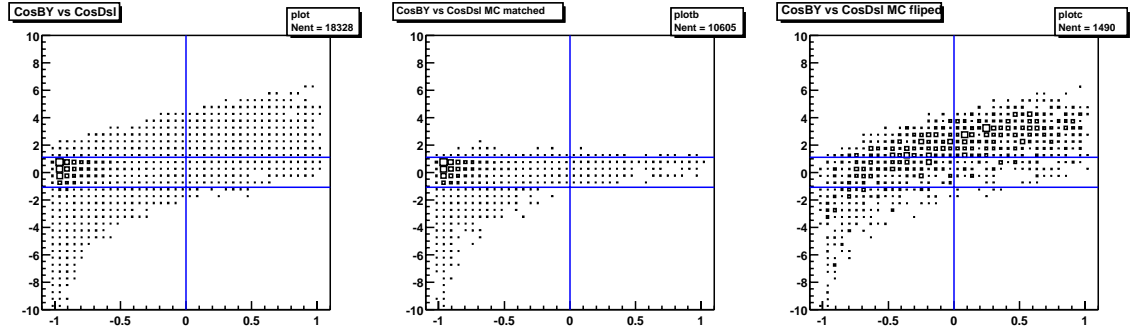
Figure 4.23: A plot of $\cos\theta_{B,D^*l}$ vs $\cos\theta_{D^*l}$. The rectangle bounded by $|cos\theta_{B,D^*l}| < 1.1$ and $\cos\theta_{D^*l} < 0$ is the signal region. The first plot shows all events. The second shows all MC matched $D^*l\nu$ events while the third plot shows flip lepton events.
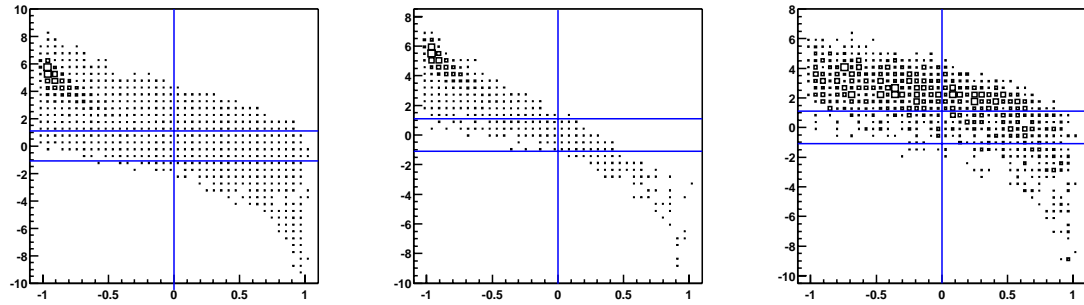
Figure 4.24: A plot of $\cos\theta_{B,D^*-l}$ vs $\cos\theta_{D^*l}$. The rectangle bounded by $|cos\theta_{B,D^*-l}| < 1.1$ and $\cos\theta_{D^*l} > 0$ is the signal region (slightly different from the previous plot!). The first plot shows all events. The second shows all MC matched $D^*l\nu$ events while the third plot shows flip lepton events.
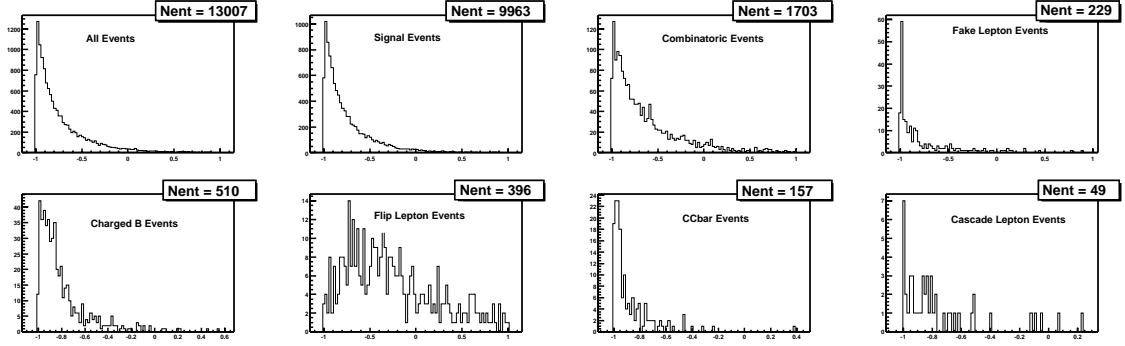
Figure 4.25: This plot shows the $\cos\theta_{D^*l}$ distribution broken down by signal and all background categories. All DstarlnuUser cuts are applied plus the D^* - D^0 mass cut (but no $\cos\theta_{D^*l}$ cut). There are two interesting things to note, first, the uncorrelated backgrounds (flip leptons) are much more uniform across $\cos\theta_{D^*l}$ than any other category (it would be almost flat if the $\cos\theta_{B,D^*l}$ cut wasn't made) and secondly, the number of events gives a good idea of the relative importance of each background. The numbers roughly correspond to the numbers in the signal sample except for the flip leptons, which lose almost a third of its events due to the $\cos\theta_{D^*l} < 0$ cut.

4.7.4 Opposite side cuts

Once a complete $D^*l\nu$ candidate has been formed and selected as a good candidate, the opposite side vertex is determined (this process is discussed in detail in section 5). This result leads to a determination of the time difference between B^0 decays, Δt, and an associated per-event error, $\sigma_{\Delta t}$. The cuts associated with these quantities are:

- $|\Delta t| < 18ps.$

- $\sigma_{\Delta t} < 2.4$ ps.

These cuts are examined and tightened further in section 5.

4.8   Cut efficiency

This section provides a running count of how many events are removed, both signal and background, as a result of various cuts applied to the no tagbit sample. The cuts considered are the D^0 cuts, D^* cuts and the $D^* l \nu$ cuts. Tables 4.10 and 4.11 show the statistics for each level of applied cuts.

The branching fractions used by BaBar Monte Carlo do not always match the PDG values, particularly for modes with branching fractions given only as upper limits. The Monte Carlo generator requires that all the branching fractions add to one, thus some fractions are scaled to meet this requirement. These fractions can change from release to release, table 4.8 shows the exact numbers used for this batch of Monte Carlo data.

Decay Mode	Branching fraction
$\pi^0 \to \gamma\gamma$	98.80 %
$K_S \to \pi^+ \pi^-$	68.61 %
$D^0 \to K\pi$	3.83 %
$D^0 \to K\pi\pi\pi$	7.49 %
$D^0 \to K\pi\pi^0$	13.9 %
$D^0 \to K_S \pi\pi$	2.7 %
$D^* \to D^0 \pi$	68.3 %
$B \to D^* l \nu$	5.6 %

Table 4.8: Branching fractions used by EvtGen, release 8.8.0c-physics-1.

The numbers given in table 4.8 indicate that 2.06% of the generic B^0's will decay into one of the four decay chains used for this analysis. The no tagbit sample has exactly 512,000 generic $B^0 \overline{B}^0$ events, which should yield 21,075±145 MC matched events (an extra factor of 2 is included since either B^0 could decay as $D^* l\nu$). The actual number found in the no tagbit sample is 4719 (22.4% of total yield). The difference is due to events falling outside detector acceptance, track inefficiency (especially the slow pion)

and the unavoidable lepton momentum cut at 0.8 MeV. The efficiencies calculated in

tables 4.9, 4.10 and 4.11

Decay Mode	MC truth expected	Actual events	acceptance fraction
$K\pi$	3000	1177	39.2%
$K\pi\pi\pi$	5867	1378	23.5%
$K\pi\pi^0$	10757	1827	17.0%
$K_S\pi\pi$	1451	337	23.2%
Total	21075	4719	22.4%

Table 4.9: Calculating the product of detector acceptance, track efficiency and Composition Tools cuts efficiencies, by decay mode.

Decay Mode	No Cuts				D^0 Cuts			
	signal events	back events	sig eff	S/B	signal events	back events	sig eff	S/B
$K\pi$	1177	26,653	39.2%	0.044	1050	14,587	35.0%	0.072
$K\pi\pi\pi$	1378	406,127	23.5%	0.003	1010	58,176	17.2%	0.017
$K\pi\pi^0$	1827	604,501	17.0%	0.003	823	30,137	7.7%	0.027
$K_S\pi\pi$	337	317,065	23.2%	0.001	61	4,455	4.2%	0.014
Total	4719	1,354,346	22.4%	0.003	2944	106,355	14.0%	0.028
	D^* Cuts				$D^*l\nu$ Cuts			
$K\pi$	1041	10,738	34.7%	0.097	373	170	12.4%	2.194
$K\pi\pi\pi$	1004	48,665	17.1%	0.021	383	842	6.5%	0.455
$K\pi\pi^0$	813	24,056	7.6%	0.034	314	492	2.9%	0.638
$K_S\pi\pi$	59	3,815	4.1%	0.015	28	43	1.9%	0.651
Total	2917	87,274	13.8%	0.033	1098	21,353	5.2%	0.229

Table 4.10: Summary of no cut, D^0 cut, D^* cut and $D^*l\nu$ cut efficiency. This table list the number of signal and background events remaining after the associated cuts have been applied. Note that there is no D^* - D^0 cut applied (see next table for that cut).

Figures 4.26, 4.27, and 4.28 show the amount of signal and background events

left after various cuts. Figure 4.29 shows the total number of MC matched $D^*l\nu$ events

prior to any cuts and the amount remaining after all cuts have been applied. All the plots

in this section have been made with the special 1 fb^{-1} sample.

	$D^*l\nu$ Cuts + D^*-D^0 mass cut($\pm$2.5 MeV)					
Decay Mode	MC truth expected	Within acceptance	signal events	back events	sig eff	S/B
$K\pi$	3000	1177	354	67	11.8%	5.28
$K\pi\pi\pi$	5867	1378	363	190	6.2%	1.91
$K\pi\pi^0$	10757	1827	306	199	2.8%	1.54
$K_S\pi\pi$	1451	337	27	9	1.9%	3.00
Total	21075	4719	1050	465	5.0%	2.258

Table 4.11: Summary of no cut, D^0 cut, D^* cut and $D^*l\nu$ cut efficiency. This table list the number of signal and background events remaining after the associated cuts have been applied. Note that there is no D^* - D^0 cut applied (see next table for that cut).
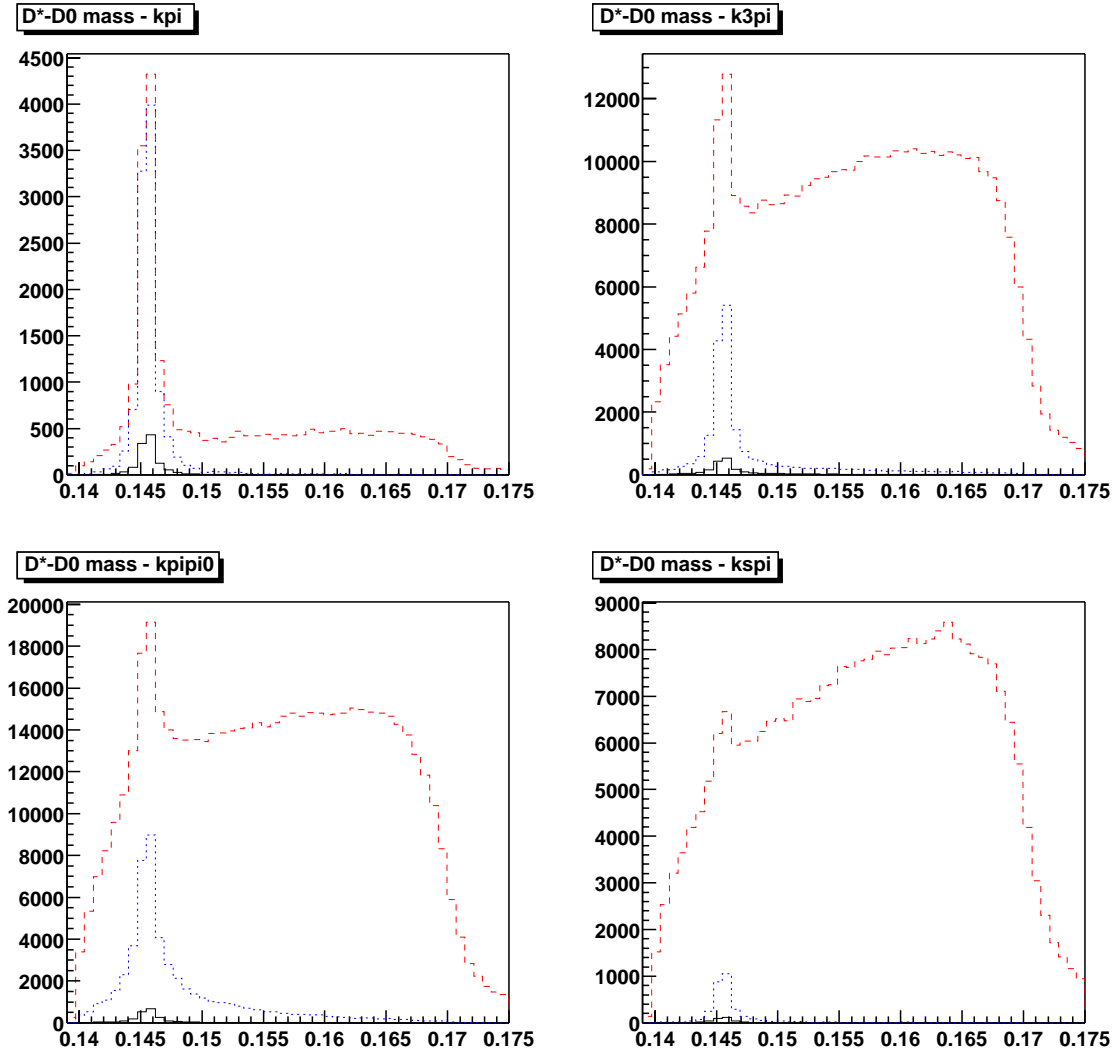
Figure 4.26: The red dash curve shows all D^*l candidate events with no cuts applied. The blue dotted curve shows all real D^*'s in this sample while the solid black curve shows all real $D^*l\nu$ events in this sample.
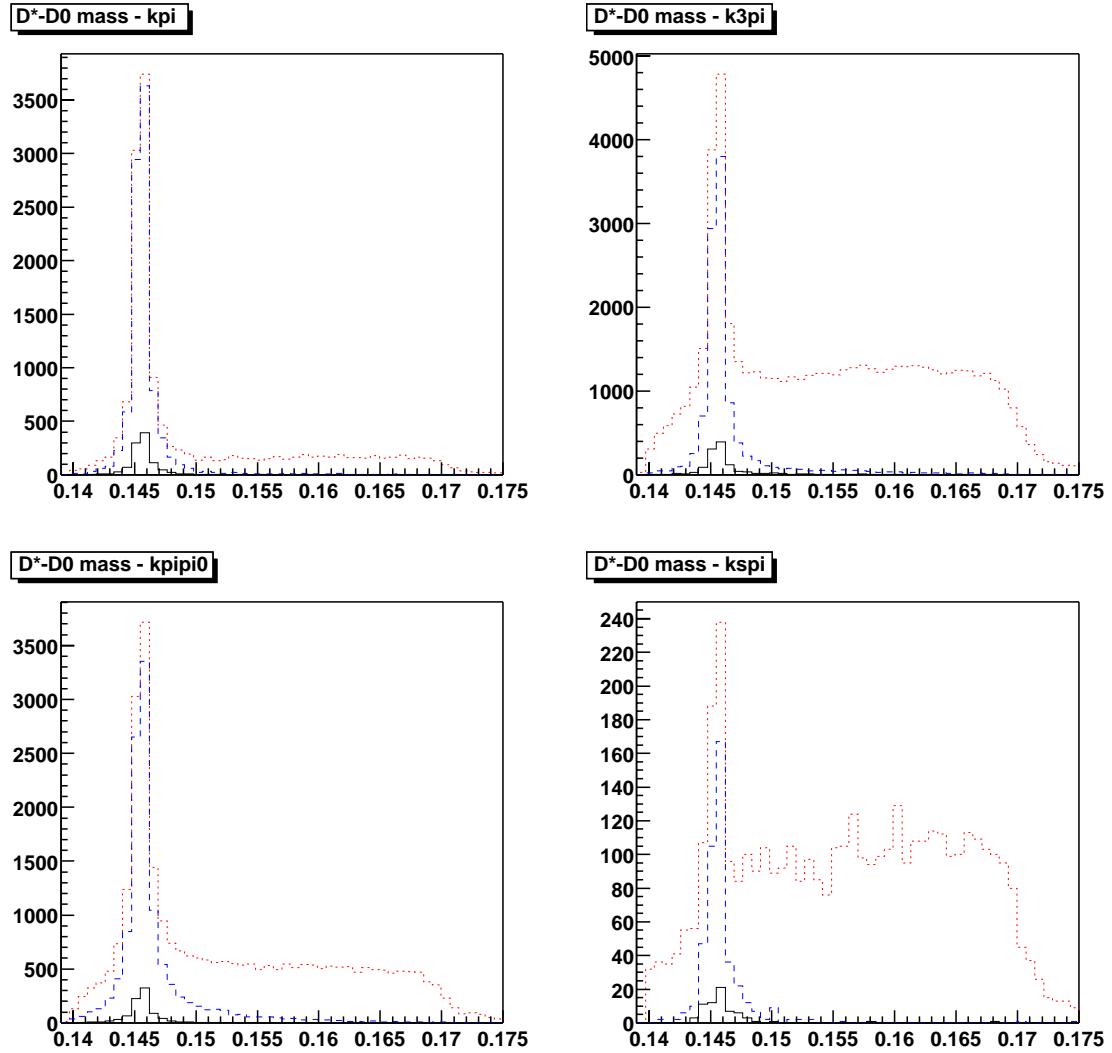
Figure 4.27: The red dash curve shows all D^*l candidate events with only D^0 cuts applied. The blue dotted curve shows all real D^*'s in this sample while the solid black curve shows all real $D^*l\nu$ events in this sample.
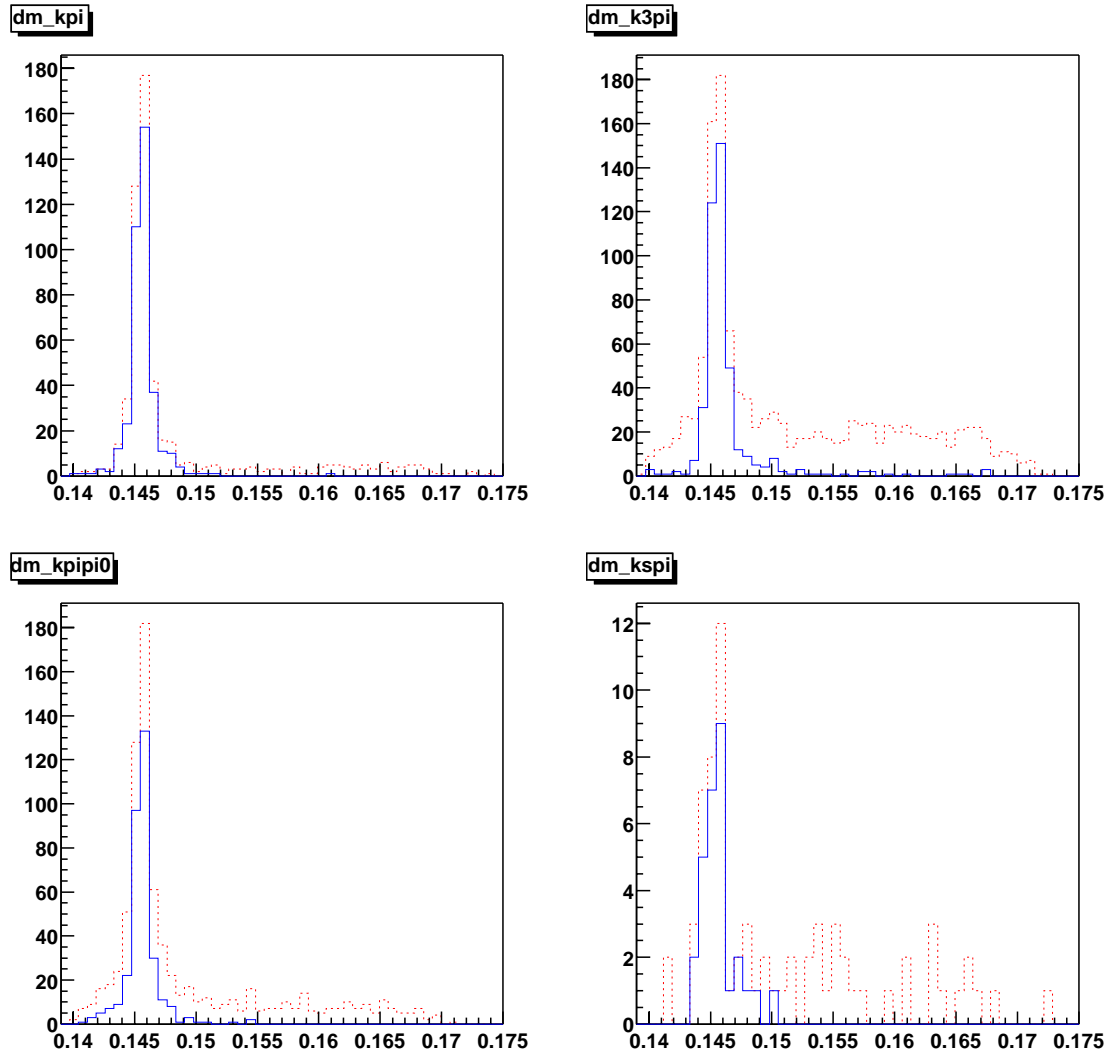
Figure 4.28: The red dash curve shows all $D^*l\nu$ candidate events with all $D^*l\nu$ cuts applied. The solid blue curve shows all real $D^*l\nu$ events in this sample.
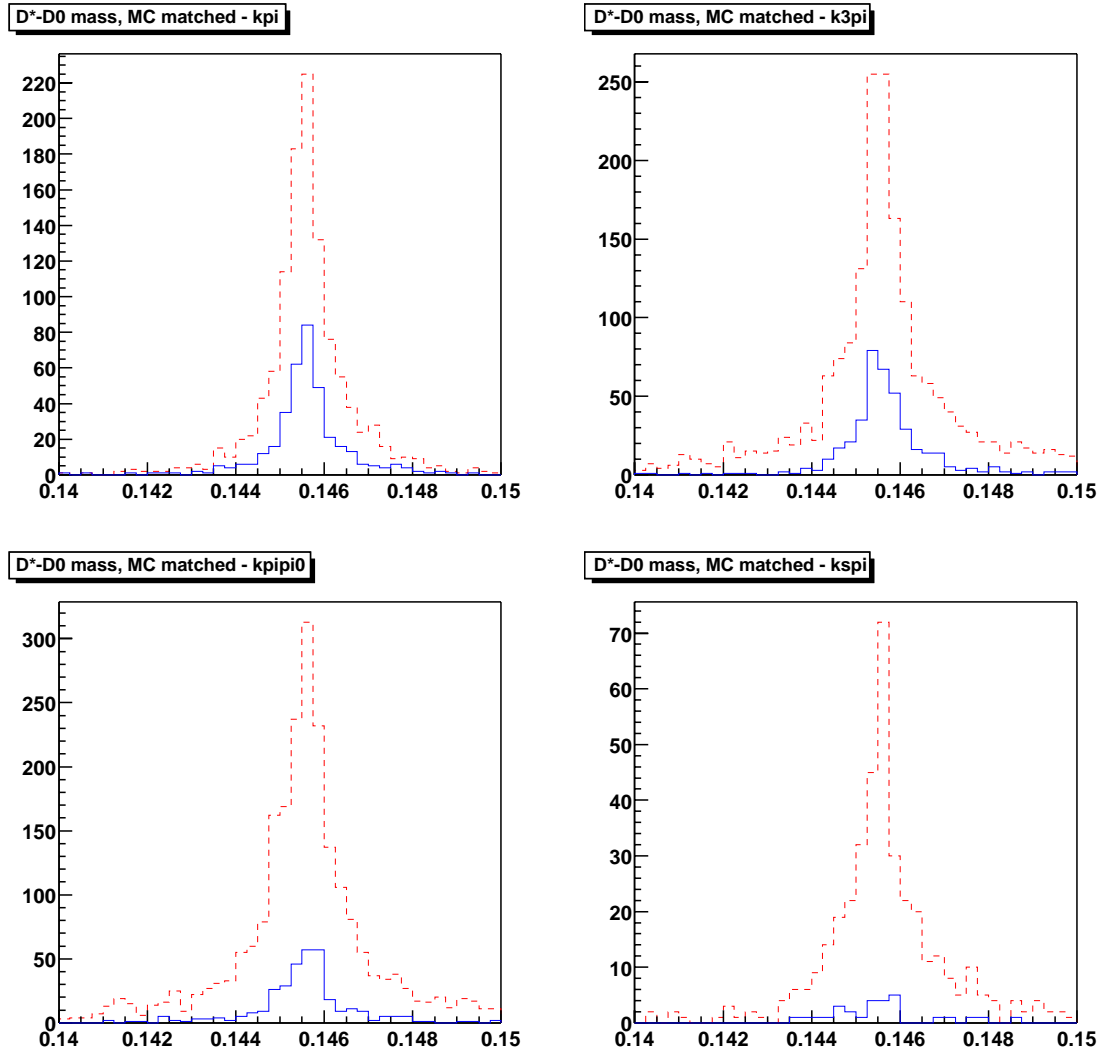
Figure 4.29: Red show all MC matched events prior to cuts, blue shows MC matched events after all cuts applied. This plot shows the efficiency for finding $D^*l\nu$ events in each decay mode.

4.9   Ntuple production

The DstarlnuUser package is used to produce HBOOK files for this analysis. The events are filtered from the "very tight" tag-bits that were written to the tag data base during skimming. For these events, the candidate lists are regenerated. The DstarlnuUserApp program was compiled in a release using the following tags on top of analysis-11:

```
CompositionUtils V00-00-16
VtxFitter V00-08-58
```

The output of each job is an HBOOK file containing an ntuple.

Each ntuple is processed using a FORTRAN selection run in PAW that applies the final selection cuts. In the case of multiple $B \to D^* l \nu$ candidates with the same D^0 decay (but different D^0 candidates), the D^0 candidate with reconstructed mass, based on simple four-vector addition, closest to the nominal mass is selected. If multiple $B \to D^* l \nu$ candidates overlap but use different D^0 decay modes, the candidates are retained. If there are multiple $B \to D^* l \nu$ candidates with the same D^0 candidate, only the first candidate in the list is kept. Table 4.2 in section 4.1 summarizes the number of candidates selected in each sample (with different D^0 decay modes combined) from each block of runs recorded during 1999–2000, before applying the final vertexing selection.

4.10   Event selection bias

The signal Monte Carlo sample described in table 4.1 is used to determine if the event selection efficiency depends on the lifetime or mixing properties of the event. The full signal Monte Carlo sample has all the DstarlnuUser cuts already applied, so any selection bias that exists prior to this stage can't be explored. The sample only uses MC

matched events and uses the Monte Carlo truth information regarding the mixing status and time difference. This ensures that any bias observed is due to event selection and not resolution effects.

The selected events are used to determine the B^0 lifetime, τ_B, the mixing frequency, Δm_d and the time-integrated mixing probability, χ_d. The lifetime and mixing frequency are determined by performing a simultaneous fit for both quantities. χ_d is determined by doing a simple counting of the number of mixed and unmixed events:

$$\chi_d = \frac{N_{mix}}{N_{mix} + N_{unmix}} \tag{4.7}$$

In Table 4.12, the difference between the measured and true values of χ_d, τ and Δm_d, after the DstarlnuUser selection criteria, and after each of the subsequent cut is shown. With all cuts applied, the biases on τ and Δm_d are -0.67σ and -0.9σ, respectively, where σ is the Monte Carlo statistical uncertainty.

This shows that the event selection is unbiased for this high statistics sample (at least any existing bias is small compared to the statistical error). Thus the event selection will not produce any significant bias in the much smaller sample using real data.

Sample	$\chi_d \times 10^{-3}$	Δm (ps^{-1})	τ (fs)	χ^2/dof unmixed	χ^2/dof mixed
	174	0.472	1548		
all cuts	-0.03 ± 1.4	-0.0012 ± 0.0013	-3.4 ± 5.1	0.930	0.796
-e only	0.92 ± 2.1	-0.0021 ± 0.0018	1.1 ± 7.6	0.644	0.632
-μ only	-0.89 ± 2.1	-0.0013 ± 0.0019	-10.8 ± 7.6	0.954	0.735

Table 4.12: Results of calculation of χ_d and simultaneous fits for Δm and τ with MC truth information. The values in the table are the difference between the measured and true values of each parameter (true value shown above double line).

Chapter 5

Decay-Time Measurement

Decay time measurement is critically important for the oscillation frequency measurement. This is the second step in the Δm_d measurement as outlined in chapter 3 and probably the most important in terms of controlling systematic errors which are very important for this measurement.

Figure 5.1 shows the topology of a typical event. The e^+e^- collide to form an $\Upsilon(4s)$ resonance, which then decays into a B^0 and a $\overline{B}^0$. Since the $\Upsilon(4s)$ is a scalar, the two B^0 daughters evolve in a coherent state until one B^0 decays, at which point there is exactly one B^0 and one $\overline{B}^0$. After the first B^0 decays, the second will continue to evolve in time and decay as a B^0 or $\overline{B}^0$. The two B's are boosted in the z direction because of the asymmetric collision with a $\beta\gamma \approx .56$. The boost momentum ($\approx 6 GeV$) is much larger than the momentums of the B^0's in the $\Upsilon(4s)$ rest frame ($p_B^* = \frac{1}{2}\sqrt{s - 4m_{B^0}^2} \simeq 320 MeV/c$). To a good approximation, the two B's are traveling purely in the z direction with a velocity determined by the boost. The time difference between the two decays is approximately given by

$$\Delta t = \Delta z / c\beta_z\gamma \qquad (5.1)$$

$$\Delta z = z_{reco} - z_{tag} \tag{5.2}$$

where γ is the boost factor for the $\Upsilon(4s)$ in the lab frame and β_z its velocity projected on the BaBar z-axis. A refinement to the boost approximation is covered in section 5.4.

For this measurement, the $B \to D^*l\nu$ will be fully reconstructed with the exception of the neutrino (reconstruction side B, shown in figure 5.1 in red). All the other tracks will be used to inclusively determine decay vertex and flavor of the other B (tag side B, shown in figure 5.1 in blue). The flavor of the tag side B (and mixing status of the event) is determined through a tagging procedure which is covered in the next chapter. The
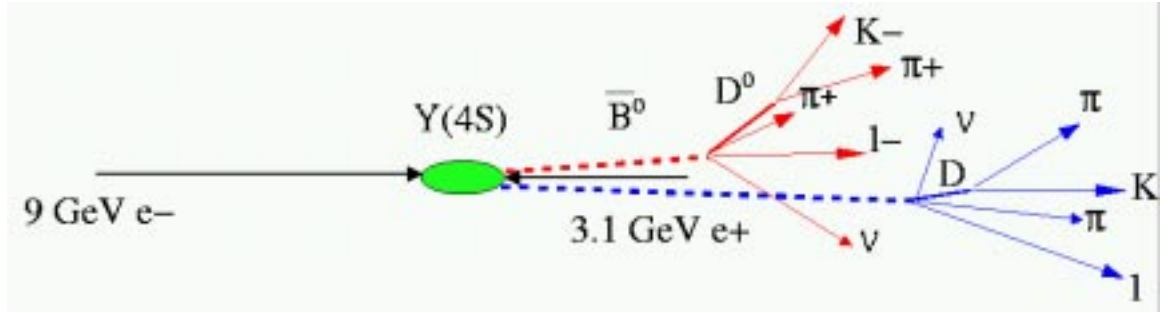
Figure 5.1: Upsilon(4S) decaying into two B's.

average separation between the two B's is given by the average lifetime times velocity:

$$< \Delta z > \ = \ < \beta\gamma c\tau > \ \approx 250\mu m \tag{5.3}$$

The importance of the boost can be seen from the fact that without an asymmetric collision, the boost provided just by the $\Upsilon(4s)$ decay into B's is

$$\beta\gamma = \frac{p_B E_B}{m_B^2} = \frac{E_B\sqrt{E_B^2 - m_B^2}}{m_B^2} \approx 0.0647 \tag{5.4}$$

Each B will travel $\approx 30\mu m$, for a total separation of 60 μm. The experimental resolution is about $130\mu m$ so the asymmetric boost is critical in the detector's ability to resolve the time difference between B^0 decays.

Ultimately, the ability to get a good Δt measurement depends on how good the tag side and reconstruction side vertices can be determined. Both vertices are found by using fitting algorithms developed for BaBar reconstruction software.

5.1   Fitting algorithms

Fitting algorithms are used to find the decay vertices of particles in a given event. The types of fits can be as simple as creating a composite particle from two or three daughters up to fitting an entire decay chain using multiple physics constraints. There are a number of different fitting algorithms used in BaBar, each with different advantages and disadvantages.

BaBar fitters use composite particles to deal with complex decay chains. A composite particle is an abstract object which is a combination of two or more charged tracks, neutral tracks or composite particles. The new composite particle momentum and energy are determined by simple 4-vector addition of the daughter particles and the new composite's genealogy is recorded internally. When a composite particle is formed, all the daughters are removed from the decay chain and replaced by the composite. All information about daughters is incorporated in the new composite and it's error matrices.

A common problem for vertexers is non-linearities in the fitting algorithm. This generally requires linear approximations and an iterative procedure to converge on an answer. The convergence depends on how good the initial guess is and how accurate the answer needs to be. The first guess is determined by calculating the point of closest approach (POCA) of the two daughter particles. If there are more than two daughters, the two closest daughters are used. The POCA serves as the initial starting point for the fit. If a POCA can't be determined, the BaBar coordinate system zero is used and a fit is

attempted.

The fitters used in BaBar are

- GeoKin - BaBar default fitter, most accurate, use most CPU time.

- FastVtx - Similar in size and scope to GeoKin, uses less CPU

- Add4 - Very quick, simple addition of 4-vectors. Used primarily as debugging tools or quick cut to reduce combinatorics (the vertex will be fit a second time with GeoKin or FastVtx for a better result in this case).

5.1.1 GeoKin

GeoKin is a general fitting algorithm which is based on the the least squares method using the Lagrange Multiplier technique. Physics constraints are used where ever possible to assist with the fitting. There are a number of different constraints that can be used, when possible. The constraints include:

$Vertex$: This constraint relates the common vertex of the daughters to i^{th} daughter. One equation for the bend and non-bend planes.

$Pseudo - momentum$: This is simply requiring that the momentum of the unknown mother is equal to the sum of the daughter momentums:

$$\sum_i p_{x,i} - p_x = 0, \ \sum_i p_{y,i} - p_y = 0, \ \sum_i p_{z,i} - p_z = 0 \tag{5.5}$$

$InvariantMass$: A constraint that the candidate have a mass M:

$$(E^2 - p_x^2 - p_y^2 - p_z^2)/M^2 - 1 = 0 \tag{5.6}$$

$\underline{BeamSpot}$: A χ^2 constraint which requires the vertex to come from the beam spot. The beam spot coordinates (x_BS, y_BS) are added as parameters to the fit with a covariance matrix to account for the non-zero size of the beam spot:

$$(x - x_{BS})^2 + (y - y_{BS})^2 = 0 \qquad (5.7)$$

This constraint can also be applied to a single track intersecting with the beam spot.

$\underline{Energy}$ This energy constraint is applied in a generic reference frame ($\vec{\beta}$):

$$\gamma E - \gamma \beta_x p_x - \gamma \beta_y p_y - \gamma \beta_z p_z - E^{fix} = 0 \qquad (5.8)$$

A special case of this is the beam energy constraint, used to ensure that a fully reconstructed B has an energy equal to half the beam energy. For the $\Upsilon(4s)$ rest frame, $\gamma = (E_{e^+} + E_{e^-})/m_{\Upsilon(4S)}$ and $\vec{\beta}\gamma = (\vec{p}_{e^+} + \vec{p}_{e^-})/m_{\Upsilon(4S)}$.

$$(E_{e^+} + E_{e^-})E - (\vec{p}_{e^+} + \vec{p}_{e^-})\vec{p} - m^2_{\Upsilon(4S)}/2 = 0 \qquad (5.9)$$

$\underline{Zero-lifetime}$ If a particle is known to be a resonance (cτ < 1 nm), a constraint can be applied which takes advantage of the extremely short lifetime of the resonance. Essentially, the daughters of the resonance will be constrained to have the same vertex as the mother of the resonance. In the decay $B \rightarrow D^*l\nu$, the D^* is a resonance.

$\underline{Line\ of\ flight}$ If the line of flight is known, the particle vertex can be constrained to be on the line of flight.

5.1.2   FastVtx

FastVtx is designed to minimize computer time, which makes it useful for high combinatorics vertexing. FastVtx does not work at the level of BtaCandidates like GeoKin,

and thus can't use the physics constraints enumerated above. FastVtx uses the track parameters directly and thus only considers the geometry of the vertex, not the kinematics. There are two constraints available for FastVtx, the beam spot constraint and a mass constraint. The beam spot constraint can be used during the computation of the vertex position, the mass constraint is applied after the vertex is determined due to the fact that FastVtx neglects correlations during the fitting procedure.

5.2   Reconstruction Side

The $B \rightarrow D^* l \nu$ decay has its own special algorithm designed for this decay channel. The algorithm involves the process of refitting. Refitting uses an initial fit which produces a common vertex for the D^0, lepton and soft pion. This vertex is now used as a constraint for "refitting" the soft pion to produce a new, or refitted soft pion daughter. The $B \rightarrow D^* l \nu$ vertex is determined again fitting the D^0, lepton and refitted slow pion simultaneously.
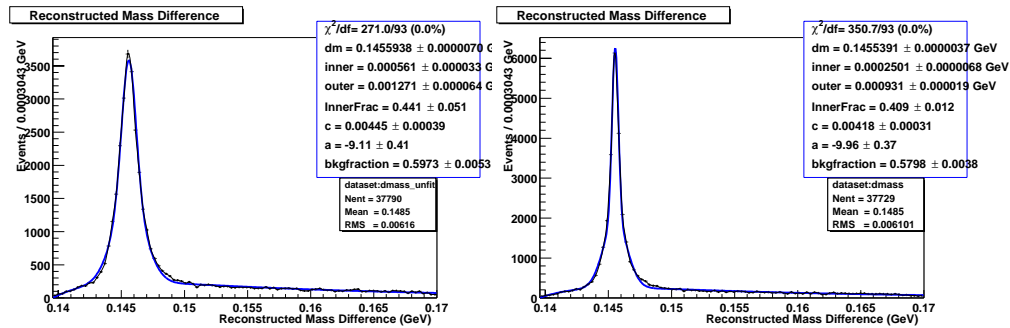
Figure 5.2: D^* - D^0 mass difference (δm) using non-refitted slow pions (left plot) and refitted slow pions (right plot).

It has been shown that refitting the soft pion from a $D^* \rightarrow D^0 \pi$ decay greatly improves the resolution on δm, as seen in figure 5.2. This results from two effects. First,

the soft pion momentum, D^0 momentum and D^0 decay length are highly correlated. Only by fitting for the D^0 vertex and D^* vertex simultaneously can the correlations be taken properly into account. Second, the D^* vertex serves as a new data point for determining the momentum of the soft pion, which helps improve the entire fit. The result is the $B \rightarrow D^* l \nu$ vertex plus improved results for the D^* vertex, D^0 vertex, and slow pion momentum.

One unexpected feature of slow pion refitting is the apparent double Gaussian shape of the signal region (both plots in figure 5.2 fit the signal with a double Gaussian, but the refitted soft pion fit clearly requires a double Gaussian). The source of the double Gaussian structure is the fact that there are two distinct populations of soft pions, ones that are found only with the SVT (SVTonly) and ones which have SVT and DCH hits (SVT+DCH). There are two effects which improve the resolution of SVT+DCH soft pion events. First, the extra hits provide a longer lever arm for measuring the particles momentum. Second, SVT+DCH soft pions must have a transverse momentum ~ 120 MeV or larger to attain the minimum 10 DCH hits to be considered a track in the DCH, so they aren't as "soft" as the other soft pions.

The end result is that the SVT+DCH events have a considerably smaller resolution (~ 0.25 MeV width) compared to the SVTonly events (~ 0.9 MeV width). Not only is the signal region different for these different types of soft pions, but the backgrounds for SVT+DCH are significantly lower as shown in figure 5.3.

The primary reason for the sharp difference between the two populations is the fact that the current reconstruction software only matches SVT tracks with established DCH tracks. In the future, SVT tracks will be projected into the DCH and search for possible hits in the DCH. This should improve the resolution of SVTonly tracks somewhat
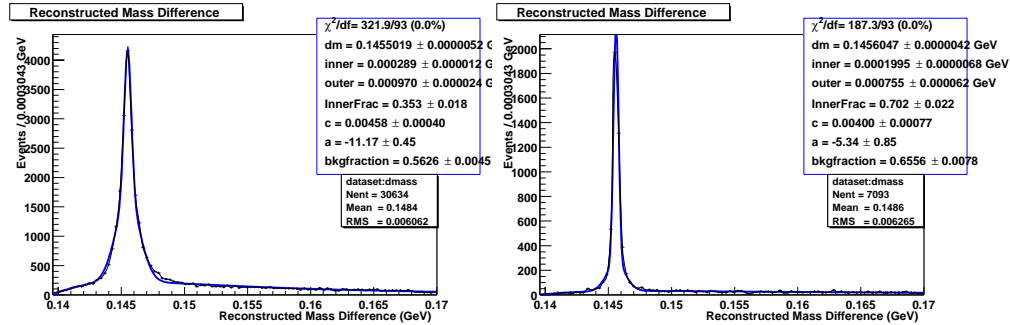
Figure 5.3: D^* - D^0 mass difference (δm) using SVTonly slow pions (left plot) and SVT+DCH- slow pions (right plot).

and blur the difference between SVTonly and SVT+DCH events.

This is one example of useful categories that have been discovered which help improve the overall description of signal and background events. The fine distinctions made here help squeeze more information out of each event. Another example of such a category is the D^0 decay mode. The signal region and background shapes are clearly different and so by considering these differences we can improve our resolution and sensitivity to the underlying physics. Several more categories will be developed in the next two chapters.

5.3 Tag Side

The tag side vertexing algorithm is called VtxTagBtaSelFit. VtxTagBtaSelFit is used in the B lifetime, B^0 oscillation frequency and $\sin(2\beta)$ measurements. VtxTagBta-SelFit uses all the particle tracks not used to create the reconstruction side candidate. The remaining tracks are fit with an inclusive technique to remove inefficiencies from requiring particular decay modes on the tag side. The biggest challenge facing the tag side fitter is to return an unbiased result. The major source of bias is long lived charm daughters on

the tag side. The tracks from charm daughters are generally displaced from the primary vertex in the positive direction due to the long charm lifetime and Upsilon(4S) boost. If these tracks are included in the determination of the tag side vertex, it will be biased towards the charm vertex. This effect can be seen on the left hand side of figure 5.4.
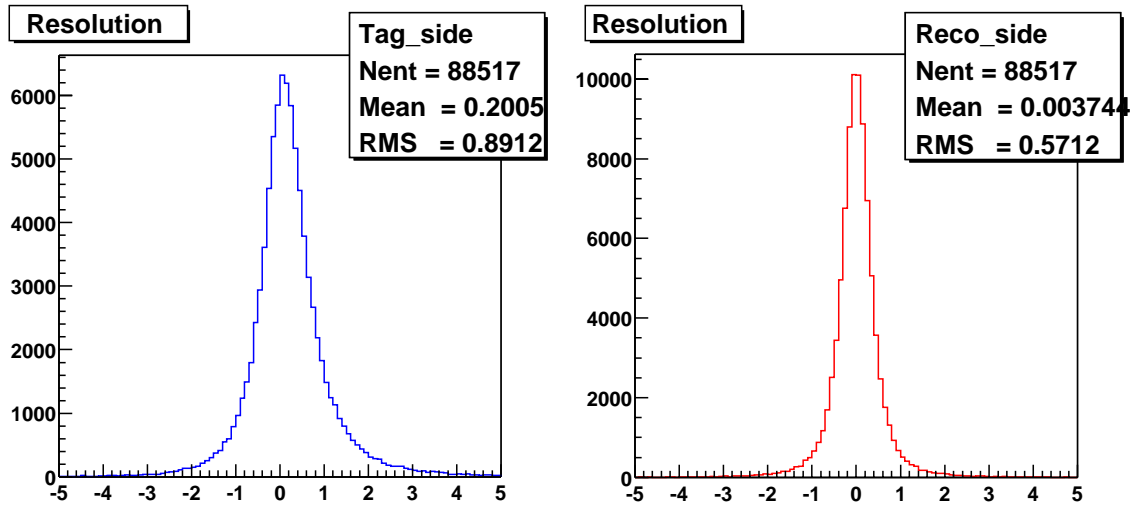
Figure 5.4: Tag side and Reconstruction side residuals (measured - truth) in ps.

The tag side vertex is determined by using the following outline:

- Reconstruct the $D^*l\nu$ candidate, create a list of all charged particle not including the tracks used on the reconstruction side.

- Remove tracks from reconstructed K_S^0 and Λ's, replacing them with the composite track.

- Fit all the remaining tracks and composites to a common vertex.

- If the χ^2 contribution for any track/composite in the fit is larger than 6 (the current value of χ^2_{min}), the track/composite with the largest χ^2 contribution is removed and

the fit is performed again. This is repeated until no tracks contribute more than 6 to the χ^2 or there is only one track remaining.

The end result is the tag side vertex, Δz and an estimate of the error on Δz (known as the "per-event-error"). The per-event-error on the difference between the $B \rightarrow D^* l \nu$ and tag side B^0 z positions takes into account the uncertainties in beam spot size, track parameters due to SVT and DCH hit resolution and multiple scattering. The error doesn't include problems in pattern recognition in tracking, problems with associating the wrong tracks with the tag side vertex or errors in the beam spot size and position. Errors associated with these possible problems will be accounted for in the resolution model of Δz in data.

5.4   Converting Δz into Δt

A number of approximations were made in equation 5.1 ($\Delta t = \Delta z / c \beta_z \gamma$) for converting Δz into Δt. The decay time difference can be determined exactly using the following formula:

$$\Delta t = t_1 - t_2 = m_{B^0} \left(\frac{r_1}{p_1} - \frac{r_2}{p_2} \right) = m_{B^0} \left(\frac{z_1}{p_1^z} - \frac{z_2}{p_2^z} \right) \tag{5.10}$$

the $\Upsilon(4s)$ decay point is defined to be at $z = 0$. The decay times t_1 and t_2 are defined in the respective B^0 rest frames. The total distance traveled by particles 1 and 2 are labelled r_1 and r_2.

The problem with determining Δt is that the decay vertex of the $\Upsilon(4s)$ is unknown. The values of t_1 and t_2 ultimately require you to know the decay vertex to get the conversion exactly right.

There are three important effects which need to be accounted for when using equation 5.1. The first is the fact that the colliding beams are not aligned with the z-axis of the detector coordinate system. The beam is rotated by $\theta \approx 20$ mrad in the x-z plane. The result is that the boost in the z-direction is reduced by $(1\text{-}\cos\theta) \times P_{\Upsilon(4s)}$ = 1.2 MeV while a boost of $\sin\theta \times P_{\Upsilon(4s)}$ = 118 MeV is created in the x direction, where $P_{\Upsilon(4s)}$ is the momentum of the Υ(4s) in the lab frame. This effect is easily accounted for by rotating the lab coordinate system before boosting. The second effect is that the Υ(4s) momentum is smeared by about 6 MeV due to the momentum spread of the colliding beams. This corresponds to a smearing of $\approx 0.1\%$ on the value of $\beta\gamma$.

The final effect is the largest, the finite momentum of the B^0's in the Υ(4s) rest frame means that the measured distance between the two B^0's is due to both the Υ(4s) boost and B^0 recoil. In the center of mass reference frame, the distance traveled in the z-direction is given by

$$\Delta z_{recoil} = c\beta^*\gamma^* cos\theta^*(t_1 + t_2) \tag{5.11}$$

where t_1 and t_2 are the B^0 lifetimes in their respective rest frame, $\beta^*\gamma^*$ is the boost of the B^0 in the Υ(4s) rest frame and θ^* is the polar angle the reconstructed B^0 makes with the Υ(4s) flight direction in the Υ(4s) rest frame. The final result for the Δz to Δt conversion is [16]

$$\Delta z = c\gamma\beta\gamma^*(t_1 - t_2) + c\gamma\gamma^*\beta^* \cos\theta^*(t_1 + t_2) \tag{5.12}$$

If the acceptance does not depend on θ^*, the average value of $cos\theta^*$ is zero. This means that Eq. 5.1 at least needs to be scaled by a factor γ^*, which is approximately 1.002.

In general we don't know $t_1 + t_2$ event by event, but if the B^0 decay angle θ^* is known (as is the case for fully reconstructed B^0's), the estimate of event-by-event Δt

can be improved using the expectation value of $t_1 + t_2$. Since both t_1 and t_2 are positive, the minimum value of $t_1 + t_2$ is $|\Delta t|$. By integrating $t_1 + t_2$ from $|\Delta t|$ to infinity, we get $\langle t_1 + t_2 \rangle = \tau_{B^0} + |\Delta t|$. Substituting this correction in Eq. 5.12, we get

$$\frac{\Delta z}{\gamma \beta \gamma^* c} = \Delta t + \frac{\beta^* \cos \theta^*}{\beta} (\tau_{B^0} + |\Delta t|) \tag{5.13}$$

For the $B \to D^* l \nu$ measurement, we can't measure the direction of the reconstructed B^0 so the best that can be done is to estimate the effect of not including this correction on the resolution of Δt. Since the correction averages to zero, the estimate will be unbiased but have a larger RMS width. The RMS contribution of the correction term is

$$c \gamma \gamma^* \beta^* \sqrt{\langle \cos^2 \theta^* \rangle \langle (t_1 + t_2)^2 \rangle} \tag{5.14}$$

For a given Δt, the expectation value of $(t_1 + t_2)^2$ depends on $|\Delta t|$. We can integrate $(t_1 + t_2)^2$ from $|\Delta t|$ to infinity and get

$$\langle (t_1 + t_2)^2 \rangle |_{\Delta t} = 2\tau_{B^0}^2 + 2\tau_{B^0}|\Delta t| + |\Delta t|^2 = \tau_{B^0}^2 + (\tau_{B^0} + |\Delta t|)^2 \tag{5.15}$$

Thus the smearing depends on $|\Delta t|$. Substituting in the values for all parameters except τ_{B^0} the RMS smearing is given by

$$0.0557 \sqrt{\tau_{B^0}^2 + (\tau_{B^0} + |\Delta t|)^2} \tag{5.16}$$

This effect can be demonstrated by plotting the RMS of $\Delta z_{\text{truth}}/\gamma \beta \gamma^* c - \Delta t_{\text{truth}}$ in bins of $|\Delta t_{\text{truth}}|$. Figure 5.5 shows this plot using 80k signal MC events that pass all our final signal cuts and 100k events directly from the event generator, and a fit to function $p_0 \cdot \sqrt{p_1^2 + (p_1 + |\Delta t|)^2}$.
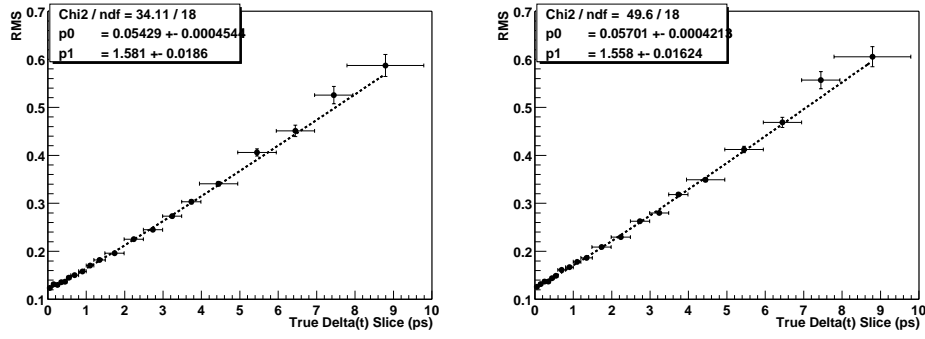
Figure 5.5: RMS of $\Delta z_{\text{truth}}/\gamma\beta\gamma^*c - \Delta t_{\text{truth}}$ residual in bins of $|\Delta t_{\text{truth}}|$ using 80k signal MC (left) and 100k events from event generator (right), and a fit to a function $p_0 \cdot \sqrt{p_1^2 + (p_1 + |\Delta t|)^2}$.

5.5   A Monte Carlo study of outlier events

Every $D^*l\nu$ candidate is assigned a value of Δt and an estimated error on Δt, $\sigma_{\Delta t}$. Occasionally, about 1% of the time, an event will be very poorly described. As an operational definition for simulated Monte Carlo events, an outlier is defined as an event where the difference between the measured value and actual value of Δt exceeds $5 \times \sigma_{\Delta t}$. The probability of an event being greater than 5 σ with a correct computation of per-event-error is 5.733×10^{-7}, quite a bit smaller than the 1% of events seen in simulated data.

Outliers come from several different sources. Individual Monte Carlo events which qualified as outliers were studied in detail and determined that outliers come from:

- The $D^*l\nu$ candidate correctly reconstructed, but the tag side B^0 vertex is misreconstructed either because a secondary charm vertex is found or a mixture of primary and secondary tracks is used to reconstruct the tag side vertex

- Flip leptons: The D^* candidate is correctly reconstructed, but the lepton actually

came from the tag side B^0.

- Reconstruction side B^0 misreconstructed: The event is completely misreconstructed, a mix of charged tracks from both B^0ś is used in each vertex.

The last two items will tend to decrease Δt on average, but they do not introduce a bias in the value of Δt. The first item leads to a negative bias in Δt since $\Delta t = t_{reco} - t_{tag}$ and t_{tag} is biased in the positive direction (as seen in figure 5.4).
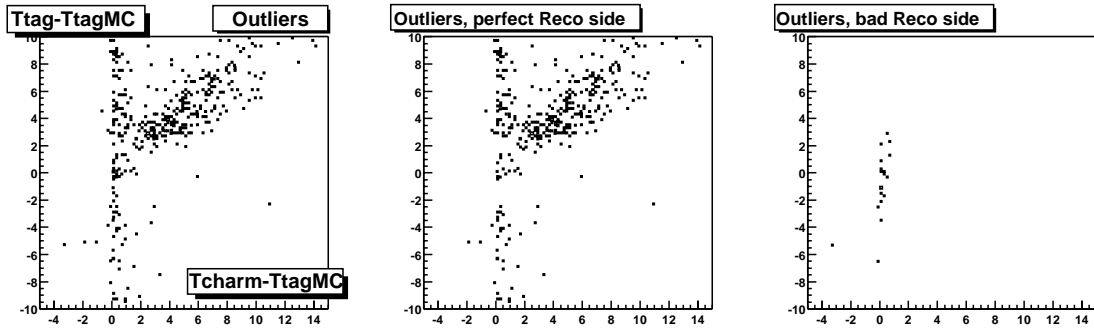
Figure 5.6: Scatter plots of outliers in signal Monte Carlo events in which the tagging B decayed to a charm particle. In each plot, the horizontal axis corresponds to the flight length of the charm particle in the z direction divided by $\beta\gamma c$. The vertical axis is the residual for the reconstructed tagging B decay time: $t_{\text{tag}}^{\text{meas}} - t_{\text{tag}}^{\text{true}}$. All units are ps.

In figure 5.6, the sources of different types of outliers are examined. The three plots contain signal Monte Carlo events with $|\delta\Delta t/\sigma_{\Delta t}| > 5$ where one of the tagging B^0's daughters eventually decays into a long lived charm daughter ($D^0, D^*, D_S, \Lambda_c, \Xi_c$, or an Ω_c). In the Monte Carlo study, this happened 96% of the time for all events in the sample and 99% of the time for outlier events (so figure 5.6 essentially uses all events in the sample). The vertical axis is the Δt residual, $t_{tag} - t_{truth}$, while the horizontal axis is the difference between the charm vertex position and the tag side B^0 vertex position converted into ps using the same conversion as equation 5.1. The left-hand plot shows all outlier events, the middle shows all outlier events with a properly reconstructed $D^*l\nu$

candidate while the right hand-side shows outliers where the reconstruction side B^0 was misreconstructed according to Monte Carlo truth information. The diagonal band shows events where the charm vertex was picked up over the B^0 vertex by the tag side vertexer. The vertical band contains events which show up as outliers but are not dominated by a charm vertex. The plots give a good idea of the relative amounts of outliers due to some sources. The number of outliers due to the reconstruction side B^0 being misreconstructed is very small, the dominant source is charm daughters used in the tag side vertex while the rest is made up of flip leptons, non-charm secondary tracks or just bad fits to the available tracks.
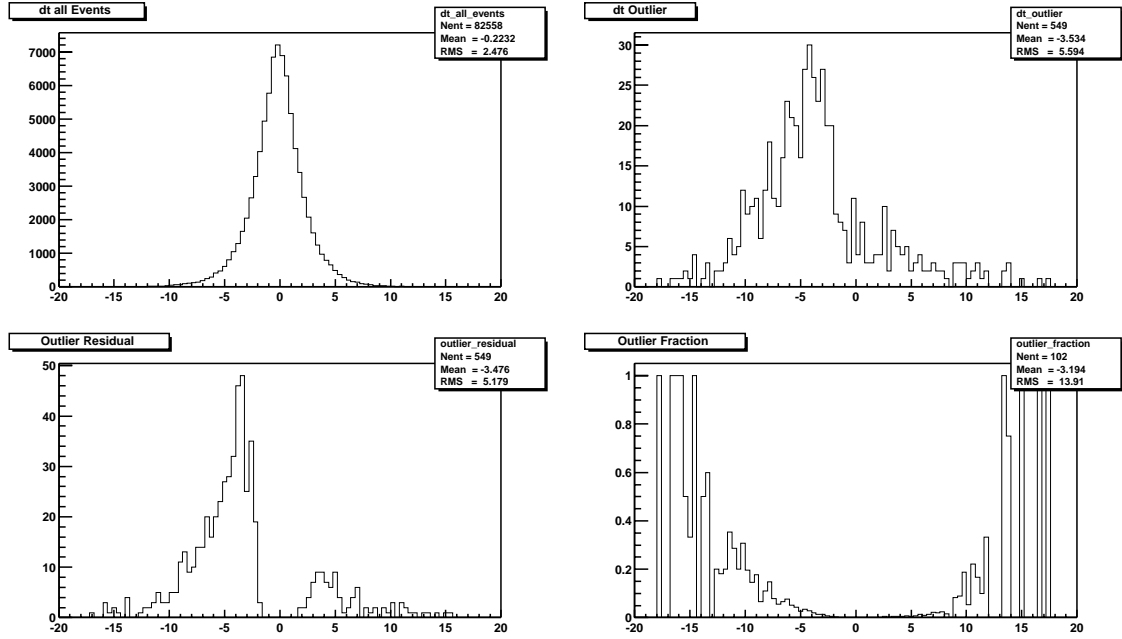
Figure 5.7: Distribution of Δt for all signal Monte Carlo $B \to D^* l\nu$ candidate events (top left) and outliers (top right). Bottom left plot shows the distribution of outlier residuals $\delta\Delta t = \Delta t_{\mathrm{reco}} - \Delta t_{\mathrm{true}}$. Bottom right plot shows the outlier fraction as a function of Δt.

A number of variables were studied to determine if they had any ability to reject outliers. The variables include Δt, the number of tracks used to create the tag side vertex and the tagging category for the event. Figure 5.7 shows the Δt distributions for all

events (first plot) and outlier events (second plot). The Δt distribution for outlier events is much broader than non-outlier events as expected. The third plot shows the distribution of residuals, $\Delta t_{measured} - \Delta t_{truth}$ for outlier events and the final plot show the percentage of outliers in different bins of Δt. One significant point is that the larger values of $|\Delta t|$ have a much larger fraction of outliers than the smaller ones. The cut on $|\Delta t|$ remains at 18 ps, but this could be tighten in the future.
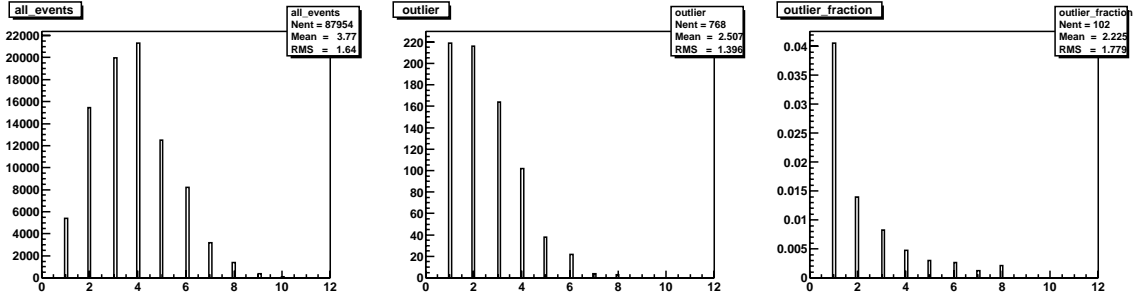
Figure 5.8: Distributions of the number of tracks used in the tagging B vertex for signal Monte Carlo. The first plot is for all events, and the second for outliers. The third plot shows the outlier fraction in each bin.

Figure 5.8 shows the number of outliers in our sample as a function of the number of tracks used to create the tag side vertex. The first plot shows all events in bins of tag side tracks, the second is number of outliers in bins of tag side tracks while the final plot show the percentage of outliers in each bin. The fraction of outliers is largest in the bins with a small number of tracks used for the vertex. This is expected since the larger the number of tracks used, the lower the chance a single bad track could throw off the result. The single track category contains about 6% of all events, but 28% of all the outliers, for this reason the single track vertices are removed from this measurement. This represents a trade off between reducing the potential systematic uncertainties due to outliers vs. a slightly larger statistical uncertainty.

5.6 Resolution Function

A resolution model describes what the residual distribution ($\Delta t_{measured} - \Delta t_{truth}$) should be for an event with measured $\Delta t_{measured}$ and a calculated per-event-error $\sigma_{\Delta t}$. The resolution model should only depend on the residuals and calculated error, not the value of Δt itself. If the per-event-error is calculated correctly and Δt is unbiased, the pull distribution ($\delta\Delta t / \sigma_{\Delta t}$) will be a unit Gaussian centered at zero. The resolution function will be given by:

$$R(\delta\Delta t, \sigma_{\Delta t}) = G(\delta\Delta t, 0, \sigma_{\Delta t}) \tag{5.17}$$

where the Gaussian function is defined as

$$G(x; x_0, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp(-(x - x_0)^2/(2\sigma^2)) \tag{5.18}$$

which has no free parameters.
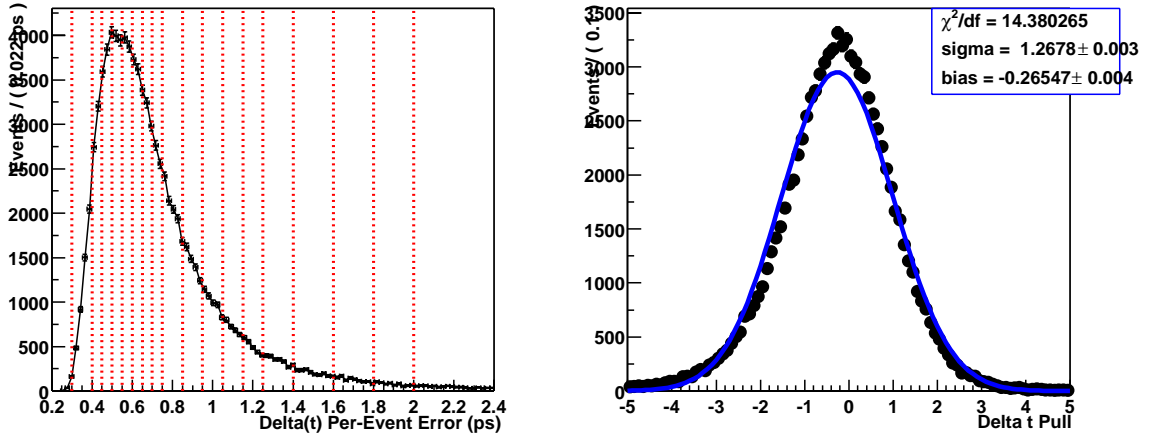
Figure 5.9: The events in these plots are correctly reconstructed $D^*l\nu$ signal Monte Carlo events with outliers ($|\delta\Delta t/\sigma_{\Delta t}| > 5$) removed. Left: Distributions of the per-event-error broken into bins of roughly equal number of events (used for investigating the scaling of bias and width vs $\sigma_{\Delta t}$. Right: Pull distribution for sample, the superimposed curve is a single Gaussian fit. The poor fit to a single Gaussian leads to the use of the GExp or double Gaussian models for the core events (all events excluding outliers).

The single fit to the pull distribution of the "core" region (outliers ($|\delta\Delta t/\sigma_{\Delta t}|>5$) are removed) on the right hand side of figure 5.9 shows three distinct problems with the single Gaussian model. The first is that the RMS is 1.266 instead of 1, this indicates that the per-event-errors are under-estimated by $\approx 27\%$. This is accounted for by including a scale factor in the resolution model to multiply the per-event-error to obtain a unit width. The second problem is that Δt in biased, which will required a second parameter to be added to the resolution model to account for this problem. The final problem is the hardest to deal with, the shape of the pull distribution is not exactly Gaussian. In order to account for this, two new resolution models have been developed which are able to better fit the pull and residual distributions of the core region.

The first model is a double Gaussian where the bias and widths are scale factors multiplying the per-event-error:

$$R_{G+G}(\delta\Delta t, \sigma_{\Delta t}; b_1, b_2, s_1, s_2, f) = f \cdot G(\delta\Delta t; b_1\sigma_{\Delta t}, s_1\sigma_{\Delta t}) + (1-f) \cdot G(\delta\Delta t; b_2\sigma_{\Delta t}, s_2\sigma_{\Delta t})$$

$$(5.19)$$

This model represents the simplest extension of the single Gaussian. There are now five parameters need to describe this model: two scale factors for the widths, two scale factors for the biases and one to specify the fraction of "narrow" Gaussian vs "wide" Gaussian.

The second model is called the "Gexp" model, it is the sum of a single, zero bias Gaussian with the same Gaussian convoluted with a one-sided ($\delta\Delta t < 0$) exponential $E(\delta\Delta t; \tau_{eff})$: (note that τ_{eff} is a dimensionless constant which multiplies the per-event-error to yield the effective lifetime)

$$R_{GExp}(\delta\Delta t, \sigma_{\Delta t}; \tau_{eff}, S, f) = f \cdot G(\delta\Delta t; 0, s\,\sigma_{\Delta t}) +$$

$$(1-f) \cdot \frac{1}{2\,\tau_{eff}\sigma_{\Delta t}} \int_{-\infty}^{0} du \; \exp(+u/(\tau_{eff}\,\sigma_{\Delta t})) \cdot G(u - \delta\Delta t; 0, s\,\sigma_{\Delta t}) \quad \text{(5.20)}$$

The integral in this definition can be performed analytically to give

$$R_{GExp}(\delta\Delta t, \sigma_{\Delta t}; \tau_{eff}, s, f) = f\, G(\delta\Delta t; 0, s\,\sigma_{\Delta t}) +$$

$$(1-f)\frac{1}{2\,\tau_{eff}\sigma_{\Delta t}} \exp\left(\frac{s^2 + 2\,\delta\Delta t \cdot \tau_{eff}/\sigma_{\Delta t}}{2\tau_{eff}^2}\right) erfc\left(\frac{s^2 + \delta\Delta t \cdot \tau_{eff}/\sigma_{\Delta t}}{\sqrt{2}\,s\,\tau_{eff}}\right) \text{(5.21)}$$

This model has just three parameters, a scale factor for the width, a lifetime for the one-sided exponential and a fraction between the two components being added. The motivation for this model comes from the idea that the bias in Δt is mostly produced by long lived charm daughter on the tag side. The long tail produced by the second term will ideally capture this bias and the lifetime should be correlated with the lifetimes of the charm daughters. Even though there is no bias in the Gaussian used in both terms, the one-sided lifetime term effective generates a bias through the asymmetry of the second term.

Figure 5.10 shows how the mean and RMS of the signal Monte Carlo sample in the core region vary as a function of $\sigma_{\Delta t}$. The variation of width with $\sigma_{\Delta t}$ is easy to understand, if the per-event-errors are consistently underestimated/overestimated by a certain fraction, the RMS of the pull distribution will be scaled by the same fraction. The connection between $\sigma_{\Delta t}$ and bias is very subtle. Figure 5.11 shows the tag side of a hypothetical event with a long lived charm daughter. If the tag side vertexer only uses primary tracks, the event will be unbiased. If the vertexer picks up the secondary vertex, the event will have a positive bias. The vertexer's algorithm uses the beam spot constraint and track χ^2 contributions to eliminate the secondary tracks as much as possible, but it is impossible to completely remove them so they influence the reconstructed tag positions. The reconstructed position will be a weighted average of the primary and secondary tracks.
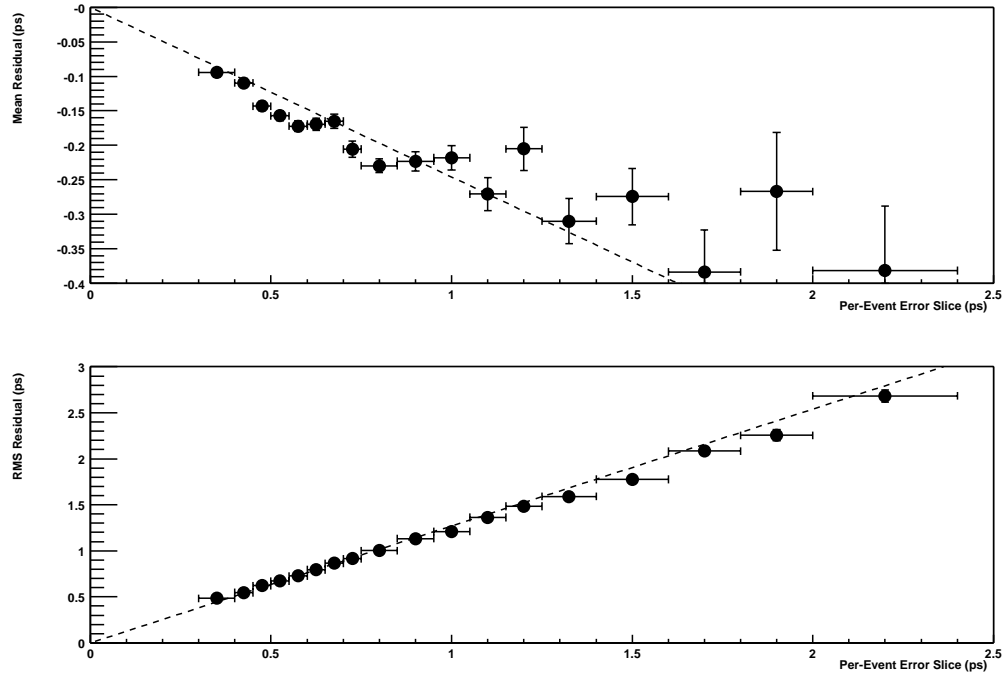
Figure 5.10: The mean and RMS of the residual distribution of correctly reconstructed $D^*l\nu$ events of signal MC in slices of $\sigma_{\Delta t}$. The dashed lines represent slopes of -0.25 (mean) and 1.27 (RMS) obtained from the single Gaussian fit from figure 5.9.

Figure 5.11 shows that charm daughters that come out at large angles with respect to the tag B^0 direction produce smaller biases and contribute less to the per-event-error than charm daughter coming out at small angles (perpendicular tracks produce smaller vertex errors than parallel tracks). This mechanism is the primary way in which residual bias correlates to $\sigma_{\Delta t}$, as shown in a detailed study by Jan Stark [28]. Toy Monte Carlo studies testing this hypothesis have reproduced the same general shape observed in signal and generic Monte Carlo samples.

Figure 5.12 shows how the parameters of R_{G+G} and R_{Gexp} vary as a function of $\sigma_{\Delta t}$. The linearity of the bias and width scale factors starts to break down as $\sigma_{\Delta t}$ gets large, which motivates a tighter cut on $\sigma_{\Delta t}$ for signal sample events. The cut of $\sigma_{\Delta t} <$ 1.8ps provides a sample inside the linear region of for most parameters while minimizing
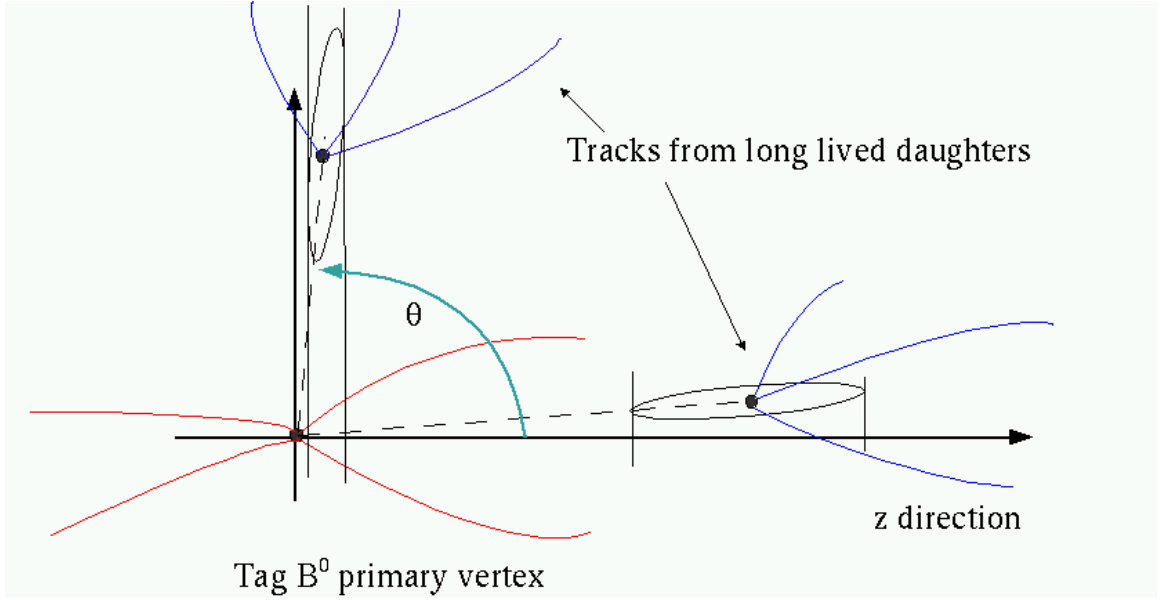
Figure 5.11: This figure shows the correlation between the bias due to long lived daughter tracks in the tag side vertex and their contribution to the calculated per-event-error.

the amount of data lost to the cut. The events which are thrown away represent the lowest quality events in the sample (in terms of $\sigma_{\Delta t}$) which also helps minimize the impact of this cut on the statistical sensitivity the Δm_d measurement.

The final step in developing the resolution model is adding the contribution of the outliers to the core model. The outliers constitute less than 1% of all events and are modeled with a single Gaussian with fixed bias and width:

$$R_{out}(\delta\Delta t, 1; b_{out}, s_{out}) = G(\delta\Delta t; b_{out}, s_{out}) \tag{5.22}$$

This leads to the full description of the two new resolution models (define $G_0 \equiv G(\delta\Delta t; 0, s\sigma_{\Delta t})$, $G_n \equiv G(\delta\Delta t; b_n\sigma_{\Delta t}, s_n\sigma_{\Delta t})$, $\mathbf{E} = E(\delta\Delta t; \tau_{eff})$)

$$R_{G+G+G}(\delta\Delta t, \sigma_{\Delta t}; b_1, b_2, b_{out}, s_1, s_2, s_{out}, f, f_{out}) = f \cdot G_1 + (1 - f - f_{out}) \cdot G_2 + f_{out} \cdot G_{out} \tag{5.23}$$

$$R_{Gexp+G}(\delta\Delta t, \sigma_{\Delta t}; \tau_{eff}, s, f, b_{out}, s_{out}, f_{out}) = f \cdot G_0 + (1 - f - f_{out}) \cdot G_0 \otimes E + f_{out} \cdot G_{out} \tag{5.24}$$
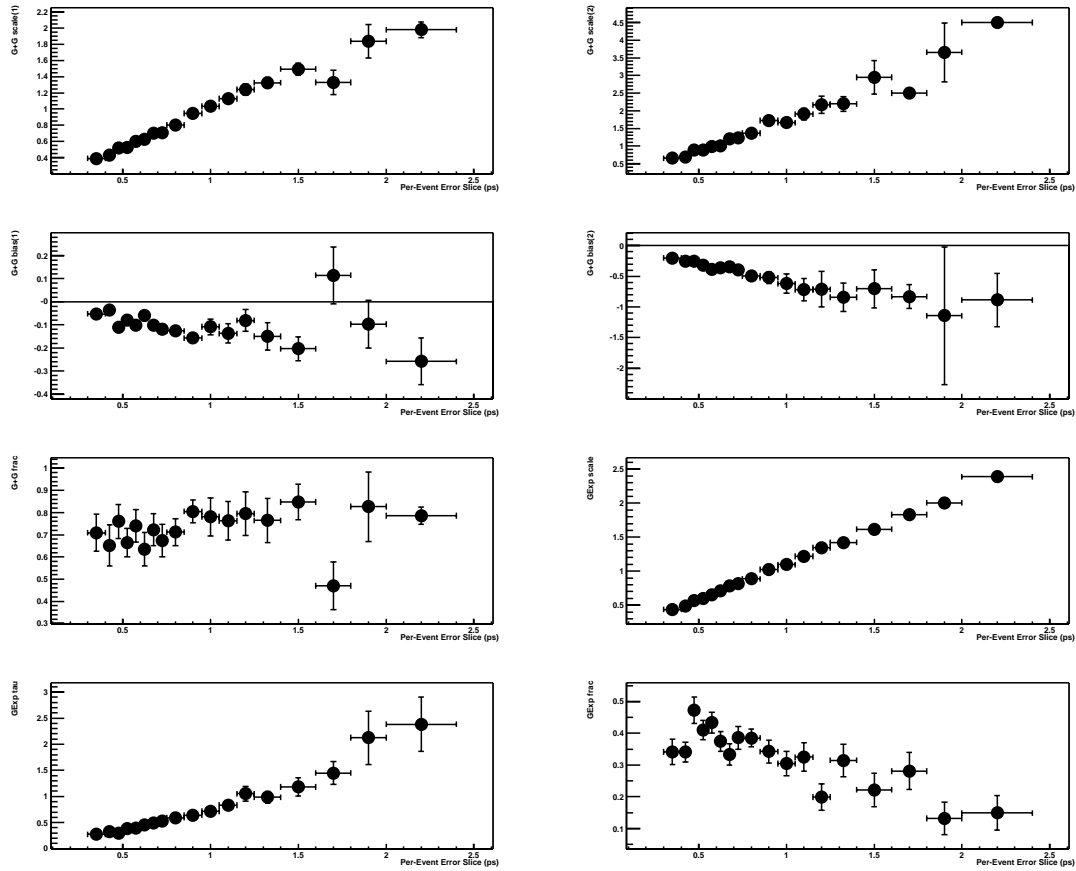
Figure 5.12: Parameter plots obtained from fitting the resolution models to data in slices of $\sigma_{\Delta t}$. The first five plots are R_{G+G} parameters s_1, s_2, b_1, b_2 and f respectively. The last three are R_{Gexp} parameters s, τ, and f.

5.7 Signal Monte Carlo study of the Resolution Function

This section describes the main results from a systematic study of the signal Monte Carlo sample list in table 4.1. The study helps to separate out the systematics shifts and biases due to event selection, the boost approximation and compensating effects from floating resolution parameters.

5.7.1   Models and Sample used

The Signal Monte Carlo resolution study uses signal decay modes specified in table 4.1. The decay modes are not correctly weighted by their luminosities, but this study is not used to determine parameters for the fit in data. The purpose of the study is to investigate the effect of the resolution model on fitting for the B^0 lifetime and oscillation frequency. All $D^*l\nu$ events are required to be properly reconstructed according to Monte Carlo truth and all signal region cuts including the $\sigma_{\Delta t} < 1.8$ps and single track vertex cut are required.

There are a total of three resolution models studied here, a single Gaussian model with a Gaussian outlier is added to the two models developed in section 5.6. The three models are labeled R_{G+G}, R_{G+G+G} and R_{Gexp+G}. This chapter contains the following results:

- Sample characteristics, number of events, subsample mean and RMS plus outlier fraction

- Fitted resolution parameters for all three models

- Lifetime fit using Monte Carlo truth lifetimes and z positions

- Lifetime fit using fixed and floating resolution parameters

- Summary table of lifetime fits by sub-categories

- A graphical summary of the above table

5.7.2 Results

Table 5.1 gives an overall view of the sample statistics. One interesting thing to look at is how well the per-event-errors work. Figure 5.15 shows the Δt residual RMS values broken down by sub-categories. The second plot shows the RMS quantities divided by $\sqrt{\langle \sigma_{\Delta t}^2 \rangle}$. This clearly shows that the per-event-errors are doing a good job at accounting for the differences between the various sub-categories, which will allow a single resolution model to be used across all sub-categories instead of independent models for each sub-category. Tables 5.2, 5.3 and 5.4 show the fitted values for the resolution parameters for the three different models. Figure 5.16 shows a lifetime fit using Monte Carlo truth Δt and Δz information. Figures 5.17 and 5.18 show fits for the resolution parameters of the different resolution models while figures 5.19 and 5.20 show lifetime fits with fitted and floating resolution parameters for the different resolution models.

There is a -3.4±5.1 fs difference between the fitted Monte Carlo truth lifetime (1.5446±0.0051 ps from figure 5.16) and the generated Monte Carlo lifetime (1.548 ps). The mixing frequency difference is -0.0012±0.0013 ps^{-1} between truth fitted events and the generator value of 0.472 ps^{-1}. This difference is due to event selection bias described in section 4.10. The lifetime value calculated from the Monte Carlo Δz distribution by using the boost approximation (c$\beta\gamma\gamma^*$ = 166.87 μm) yields (1.5494±0.0051) which is 4.8 fs higher than the Δt truth value. The truth Δz mixing frequency is 0.0015 ps^{-1} higher than the truth Δt mixing frequency. This difference (between truth Δt and Δz) is due to the boost approximation.

The boost approximation bias for the lifetime can be understood as follows. Imagine you have a pure double-sided exponential decay, $e^{-|x|/\tau}$, and apply a small Gaussian smearing to it (a simple model for the boost approximation resolution). Now

break this distribution into many small bins in $|x|$. In each bin, an equal number of events will migrate to larger and smaller value bins due to the smearing. This is true everywhere except for bins near x=0. Near x=0, events can migrate to larger values of $|x|$ but can't migrate to lower values due to the absolute value. Thus the smearing will cause a net shift to larger values of $|x|$.

The bias for the mixing can't be stated in such simple terms, but the underlying cause is the different Δt shapes that mixed and unmixed events have (see the first plot in figure 7.3 to see what the truth Δt shapes are for mixed and unmixed events).

The bias generated by this effect was estimated by Chih-hsiang Cheng by creating 100 toy Monte Carlo experiments by using $e^{-|x|/\tau}$ and (1-$cos^2\theta^*$) distributions and equation 5.12 to calculate Δz. The results are shown in figure 5.13. The blue vertical line shows the value obtained for the Signal Monte Carlo sample, which is consistent with the toy study. An important question to ask is does the boost approximation fully account for the dependence of the resolution models on Δt_{truth}? Figure 5.14 shows that the boost approximation is responsible for most if not all the dependence.

Table 5.5 shows the results for fitting to truth Δt, truth Δz and to just Δt with fixed and floating resolution parameters. This table shows the various biases mentioned so far, the event selection bias, boost approximation bias and bias due to the resolution models.
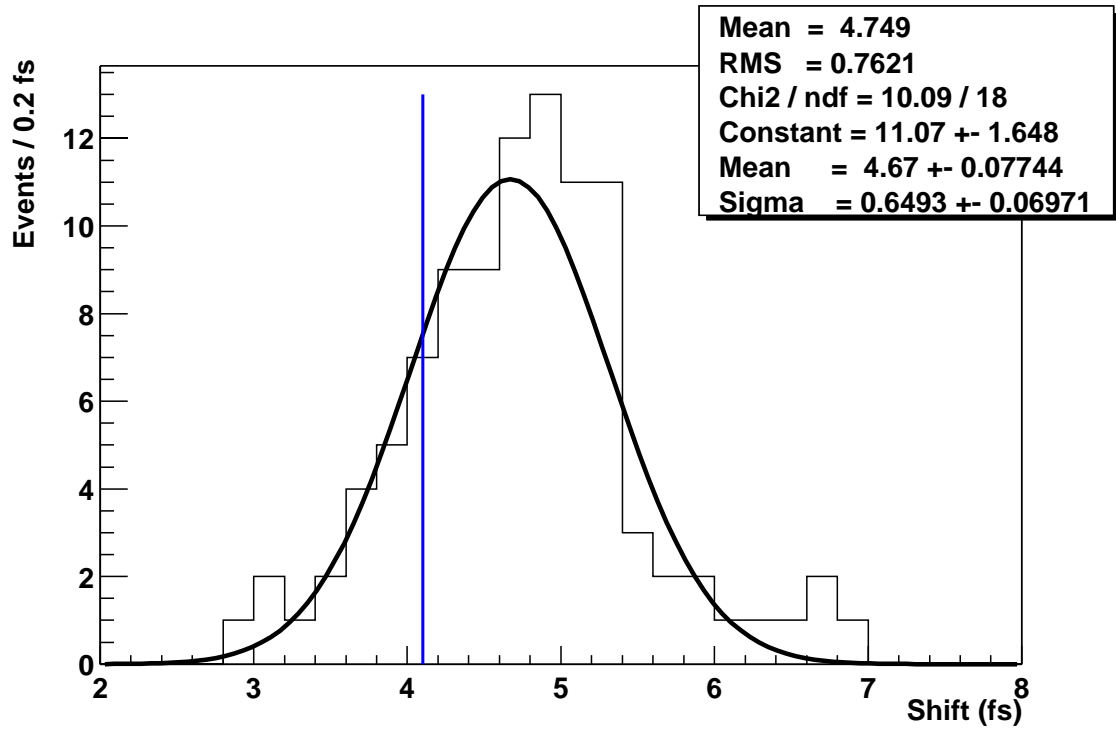
Figure 5.13: Shift of $\langle |\Delta z|/\beta\gamma\gamma^*c \rangle$ from $\langle |\Delta t| \rangle$ with 100 toy experiments. The Blue line indicates the value found in the signal Monte Carlo fit.
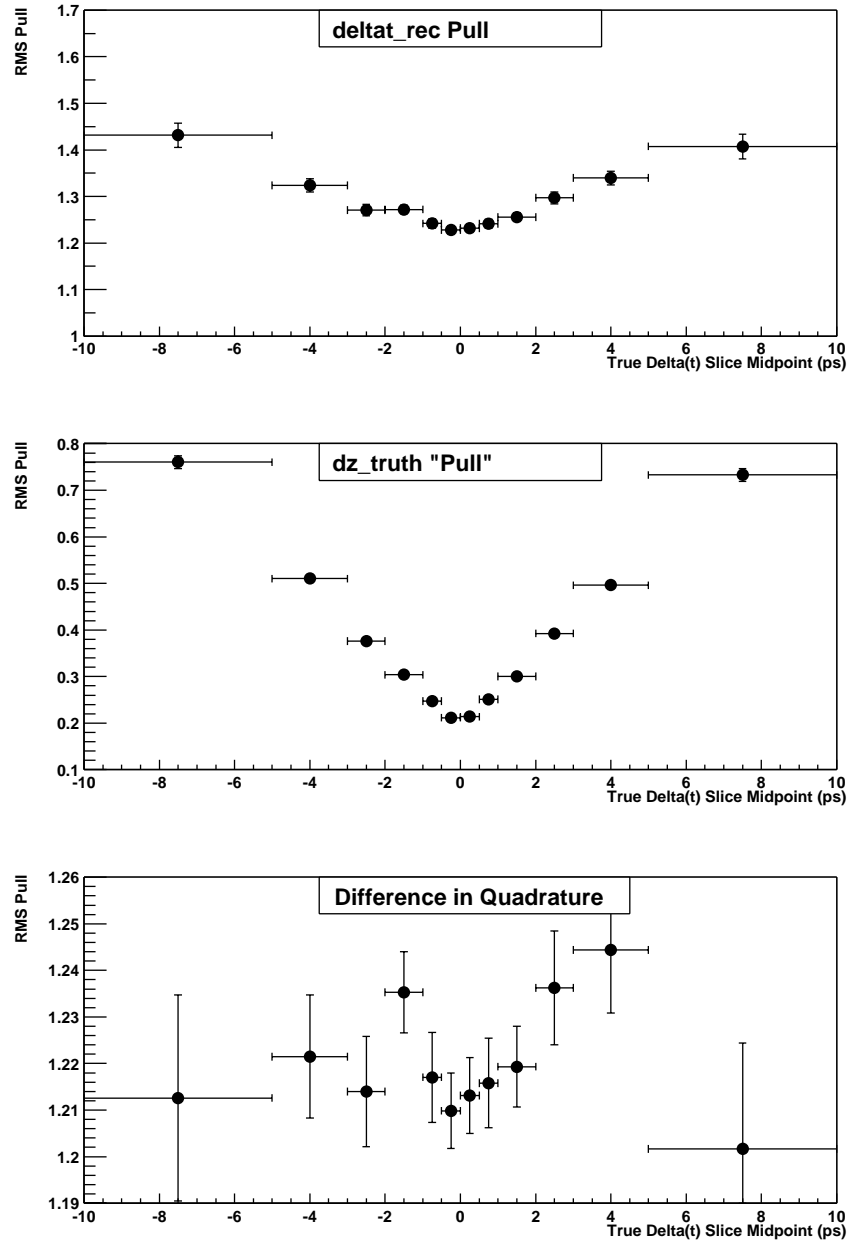
Figure 5.14: Top: the RMS of the pull (($\Delta t - \Delta t_{true})/\sigma_{\Delta t}$) in bins of Δt_{true} for signal MC. Middle: same plot with Δt replaced by $\Delta z_{\mathrm{true}}/\gamma\beta\gamma^*c$ in the calculation of the RMS of the pull. Bottom: bin-by-bin difference in quadrature between the top two histograms. Note the factor of 10 difference in the vertical scales of the top and bottom histograms.
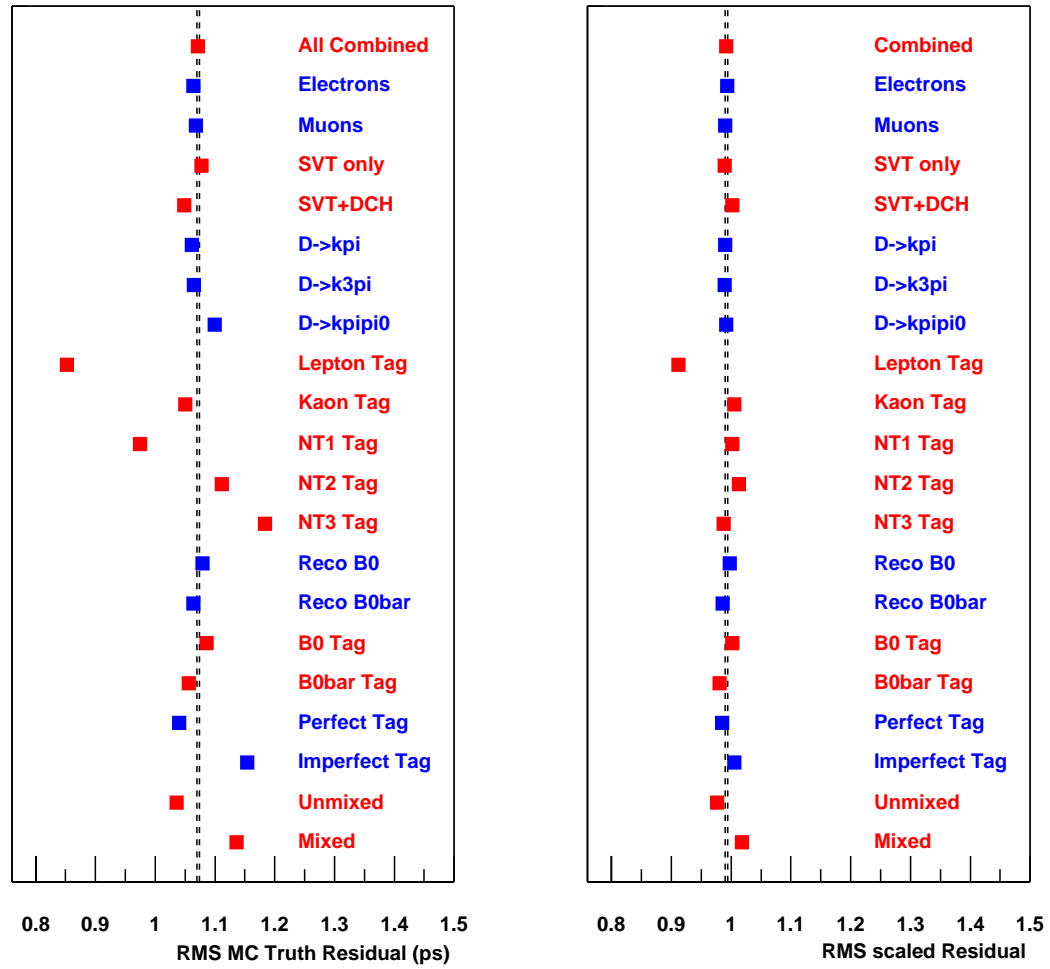
Figure 5.15: Left hand plot shows Δt residual RMS by sub-category. Right hand side is Δt residual RMS divided by $\sqrt{2\langle\sigma^2_{\Delta t}\rangle}$. This demonstrates that the estimated per-event-errors, $\sigma_{\Delta t}$, does a good jobs of accounting for differences between different categories of data.

Sample	Events	Mean $\delta\Delta t$ (ps)	RMS $\delta\Delta t$ (ps)	$\sqrt{\langle\sigma^2_{\Delta t}\rangle}$ (ps)	$f(|\mathbf{pull}| > 5)$ (%)
All Combined	**92879**	-0.2082 ± 0.004	1.072 ± 0.002	**0.764**	**0.71± 0.0275**
Electrons	**42688**	-0.2111 ± 0.005	1.064 ± 0.004	**0.758**	**0.668± 0.0394**
Muons	**40905**	-0.2015 ± 0.005	1.068 ± 0.004	**0.763**	**0.743± 0.0425**
SVT only	**74429**	-0.2060 ± 0.004	1.077 ± 0.003	**0.770**	**0.68± 0.0301**
SVT+DCH	**18450**	-0.2169 ± 0.008	1.047 ± 0.005	**0.740**	**0.829± 0.0668**
$D^0 \to K\pi$	**51838**	-0.2081 ± 0.005	1.061 ± 0.003	**0.758**	**0.706± 0.0368**
$D^0 \to K\pi\pi\pi$	**21616**	-0.2097 ± 0.007	1.064 ± 0.005	**0.761**	**0.699± 0.0566**
$D^0 \to K\pi\pi^0$	**17135**	-0.2018 ± 0.008	1.100 ± 0.006	**0.784**	**0.712± 0.0642**
$D^0 \to K_S\pi\pi$	**2290**	-0.243 ± 0.02	1.165 ± 0.02	**0.783**	**0.873± 0.194**
Lepton Tag	**12270**	-0.1057 ± 0.008	0.852 ± 0.005	**0.661**	**0.399± 0.0569**
Kaon Tag	**32509**	-0.2338 ± 0.006	1.050 ± 0.004	**0.739**	**0.689± 0.0459**
NT1 Tag	**7732**	-0.1371 ± 0.01	0.974 ± 0.008	**0.688**	**0.66± 0.0921**
NT2 Tag	**13425**	-0.2118 ± 0.010	1.112 ± 0.007	**0.776**	**0.775± 0.0757**
NT3 Tag	**26943**	-0.2425 ± 0.007	1.184 ± 0.005	**0.848**	**0.857± 0.0562**
Reco'd B^0	**47184**	-0.2105 ± 0.005	1.079 ± 0.004	**0.765**	**0.71± 0.0387**
Reco'd $\overline{B}^0$	**45695**	-0.2057 ± 0.005	1.064 ± 0.004	**0.763**	**0.709± 0.0393**
B^0Tag	**47652**	-0.2161 ± 0.005	1.086 ± 0.004	**0.767**	**0.751± 0.0396**
$\overline{B}^0$Tag	**45227**	-0.1998 ± 0.005	1.056 ± 0.004	**0.761**	**0.666± 0.0382**
Perfect Tag	**68448**	-0.1946 ± 0.004	1.040 ± 0.003	**0.747**	**0.653± 0.0308**
Imperfect Tag	**24431**	-0.2461 ± 0.007	1.154 ± 0.005	**0.812**	**0.868± 0.0593**
Unmixed	**60894**	-0.1938 ± 0.004	1.036 ± 0.003	**0.751**	**0.639± 0.0323**
Mixed	**31985**	-0.2355 ± 0.006	1.136 ± 0.004	**0.789**	**0.844± 0.0512**

Table 5.1: Statistics of the semileptonic signal Monte Carlo sample, and residual and $\sigma_{\Delta t}$ distributions.

Sample	b_1	s_1	b_2	s_2	f
All Combined	-0.23740	1.1843	-1.7593	3.880	0.9586
	±0.004	±0.004	±0.08	±0.07	±0.002
Electrons	-0.24383	1.1972	-1.917	4.22	0.9645
	±0.006	±0.006	±0.1	±0.1	±0.002
Muons	-0.23241	1.1774	-1.645	3.677	0.9550
	±0.007	±0.006	±0.1	±0.09	±0.003
SVT only	-0.23664	1.1856	-1.7287	3.900	0.9597
	±0.005	±0.004	±0.09	±0.08	±0.002
SVT+DCH	-0.24045	1.1790	-1.868	3.80	0.9540
	±0.010	±0.009	±0.2	±0.1	±0.004
$D^0 \to K\pi$	-0.23528	1.1819	-1.863	3.926	0.9603
	±0.006	±0.005	±0.1	±0.09	±0.002
$D^0 \to K\pi\pi\pi$	-0.23618	1.1817	-1.522	3.51	0.9491
	±0.009	±0.009	±0.1	±0.1	±0.004
$D^0 \to K\pi\pi^0$	-0.2373	1.1878	-1.531	3.92	0.9577
	±0.01	±0.010	±0.2	±0.2	±0.004
$D^0 \to K_S\pi\pi$	-0.2811	1.191	-3.767	5.46	0.9764
	±0.03	±0.02	±1.0	±0.6	±0.005
Reco'd B^0	-0.24565	1.1943	-1.891	4.22	0.9644
	±0.006	±0.005	±0.1	±0.1	±0.002
Reco'd $\overline{B}^0$	-0.22752	1.1715	-1.6177	3.535	0.9503
	±0.006	±0.006	±0.10	±0.08	±0.003
B^0Tag	-0.24393	1.1862	-1.944	4.064	0.9607
	±0.006	±0.005	±0.1	±0.10	±0.002
$\overline{B}^0$Tag	-0.23007	1.1809	-1.559	3.657	0.9551
	±0.006	±0.006	±0.1	±0.09	±0.003
Lepton Tag	-0.1359	1.137	-1.252	3.19	0.9616
	±0.01	±0.01	±0.2	±0.2	±0.006
Kaon Tag	-0.28711	1.2182	-1.810	3.76	0.9593
	±0.008	±0.007	±0.1	±0.1	±0.003
NT1 Tag	-0.1865	1.201	-1.882	5.39	0.9766
	±0.01	±0.01	±0.4	±0.4	±0.003
NT2 Tag	-0.2441	1.185	-1.632	4.36	0.9618
	±0.01	±0.01	±0.2	±0.2	±0.004
NT3 Tag	-0.23448	1.1528	-1.815	3.67	0.9458
	±0.008	±0.008	±0.1	±0.1	±0.004
Perfect Tag	-0.22491	1.1834	-1.7366	3.737	0.9595
	±0.005	±0.005	±0.09	±0.08	±0.002
Imperfect Tag	-0.27145	1.1852	-1.808	4.19	0.9551
	±0.008	±0.008	±0.2	±0.1	±0.003
Unmixed	-0.22461	1.1764	-1.855	3.814	0.9623
	±0.005	±0.005	±0.1	±0.08	±0.002
Mixed	-0.26235	1.2006	-1.632	3.99	0.9524
	±0.008	±0.007	±0.1	±0.1	±0.003

Table 5.2: Results of fitting a G+G resolution model to $\delta\Delta t$ and $\sigma_{\Delta t}$ for signal MC.

Sample	b_1	s_1	b_2	s_2	b_{out}	s_{out}	f	f_{out}
Combined	-0.1720	1.080	-0.8119	1.957	-2.854	6.69	0.824	0.00808
	±0.007	±0.009	±0.04	±0.05	±0.3	±0.3	±0.01	±0.0009
Electrons	-0.1735	1.085	-0.8375	1.989	-3.548	8.11	0.827	0.00594
	±0.010	±0.01	±0.06	±0.06	±0.6	±0.6	±0.02	±0.0008
Muons	-0.151	1.047	-0.6377	1.730	-2.251	5.22	0.750	0.0146
	±0.01	±0.02	±0.06	±0.07	±0.3	±0.3	±0.04	±0.002
SVT only	-0.1680	1.074	-0.7575	1.903	-2.751	6.51	0.810	0.00865
	±0.008	±0.01	±0.05	±0.05	±0.3	±0.3	±0.02	±0.0009
SVT+DCH	-0.1832	1.093	-1.005	2.138	-3.338	7.54	0.855	0.0060
	±0.01	±0.02	±0.1	±0.10	±0.9	±0.9	±0.02	±0.002
$D^0 \to K\pi$	-0.1680	1.078	-0.8210	1.942	-3.075	6.55	0.826	0.0084
	±0.01	±0.01	±0.07	±0.08	±0.4	±0.4	±0.02	±0.001
$D^0 \to K\pi\pi\pi$	-0.1795	1.096	-0.8337	1.977	-2.123	6.06	0.831	0.0090
	±0.01	±0.02	±0.10	±0.10	±0.5	±0.5	±0.03	±0.002
$D^0 \to K\pi\pi^0$	-0.1656	1.059	-0.7108	1.916	-2.382	7.14	0.789	0.0074
	±0.02	±0.02	±0.08	±0.08	±0.8	±0.7	±0.03	±0.002
$D^0 \to K_S\pi\pi$	-0.1922	1.082	-1.112	1.92	-4.000	8.1	0.86	0.0089
	±0.06	±0.07	±0.5	±0.4	±0.1	±2	±0.1	±0.005
Reco'd B^0	-0.1925	1.108	-0.9551	2.158	-3.358	8.39	0.868	0.00532
	±0.009	±0.01	±0.07	±0.08	±0.7	±0.7	±0.01	±0.0010
Reco'd $\overline{B}^0$	-0.1560	1.058	-0.7260	1.847	-2.539	5.67	0.787	0.0108
	±0.01	±0.01	±0.05	±0.06	±0.3	±0.3	±0.02	±0.001
B^0Tag	-0.1804	1.079	-0.8327	2.010	-3.659	7.15	0.830	0.0073
	±0.010	±0.01	±0.06	±0.07	±0.5	±0.5	±0.02	±0.001
$\overline{B}^0$Tag	-0.1632	1.083	-0.7992	1.908	-2.111	6.17	0.819	0.0091
	±0.01	±0.01	±0.06	±0.06	±0.4	±0.4	±0.02	±0.001
Lepton Tag	-0.1006	1.063	-0.479	1.76	-2.081	4.63	0.833	0.0094
	±0.02	±0.03	±0.1	±0.1	±0.6	±0.5	±0.05	±0.003
Kaon Tag	-0.2030	1.099	-0.8316	1.878	-2.803	6.18	0.794	0.0086
	±0.01	±0.02	±0.08	±0.08	±0.5	±0.5	±0.03	±0.002
NT1 Tag	-0.1518	1.143	-1.234	2.64	-0.38	16.9	0.930	0.0021
	±0.02	±0.02	±0.2	±0.2	±6	±7	±0.02	±0.001
NT2 Tag	-0.1981	1.105	-0.951	2.24	-2.00	8.9	0.875	0.0057
	±0.01	±0.02	±0.1	±0.1	±1	±1	±0.02	±0.002
NT3 Tag	-0.1556	1.020	-0.7329	1.852	-3.189	5.86	0.765	0.0122
	±0.01	±0.02	±0.07	±0.07	±0.5	±0.4	±0.03	±0.002
Perfect Tag	-0.1652	1.082	-0.7441	1.923	-2.984	6.12	0.822	0.0083
	±0.008	±0.01	±0.05	±0.06	±0.4	±0.3	±0.02	±0.001
Imperfect Tag	-0.1879	1.067	-0.9471	1.996	-2.430	7.89	0.814	0.0084
	±0.01	±0.02	±0.08	±0.09	±0.6	±0.7	±0.02	±0.002
Unmixed	-0.1674	1.080	-0.7889	1.958	-3.378	6.48	0.838	0.0071
	±0.008	±0.01	±0.06	±0.07	±0.4	±0.4	±0.02	±0.001
Mixed	-0.1781	1.078	-0.8396	1.935	-2.104	6.82	0.792	0.0102
	±0.01	±0.02	±0.07	±0.07	±0.5	±0.5	±0.03	±0.002

Table 5.3: Results of fitting a G+G+G resolution model to $\delta\Delta t$ and $\sigma_{\Delta t}$ for signal MC.

Sample	s	f	τ_{eff}	s_{out}	b_{out}	f_{out}
All Combined	1.0963	0.6614	0.849	5.45	-1.472	0.0159
	±0.004	±0.009	±0.02	±0.2	±0.2	±0.001
Electrons	1.1030	0.673	0.890	6.47	-1.607	0.0115
	±0.006	±0.01	±0.03	±0.3	±0.3	±0.001
Muons	1.0939	0.641	0.779	4.72	-1.428	0.0217
	±0.006	±0.02	±0.03	±0.2	±0.2	±0.002
SVT only	1.0981	0.658	0.832	5.41	-1.432	0.0161
	±0.004	±0.01	±0.02	±0.2	±0.2	±0.001
SVT+DCH	1.0897	0.674	0.922	5.66	-1.615	0.0148
	±0.009	±0.02	±0.05	±0.4	±0.4	±0.003
$D^0 \to K\pi$	1.0922	0.664	0.846	5.47	-1.658	0.0158
	±0.005	±0.01	±0.03	±0.2	±0.2	±0.002
$D^0 \to K\pi\pi\pi$	1.1044	0.651	0.835	4.89	-1.165	0.0185
	±0.008	±0.02	±0.04	±0.3	±0.3	±0.003
$D^0 \to K\pi\pi^0$	1.0996	0.664	0.853	5.49	-1.000	0.0163
	±0.009	±0.02	±0.04	±0.4	±0.4	±0.003
$D^0 \to K_S\pi\pi$	1.062	0.634	0.88	7.9	-3.0000	0.0109
	±0.02	±0.05	±0.1	±2	±0.01	±0.004
Reco'd B^0	1.0997	0.663	0.864	6.09	-1.552	0.0133
	±0.005	±0.01	±0.03	±0.3	±0.3	±0.001
Reco'd $\overline{B}^0$	1.0916	0.657	0.827	4.79	-1.385	0.0200
	±0.006	±0.01	±0.03	±0.2	±0.2	±0.002
B^0Tag	1.0925	0.657	0.856	5.70	-1.765	0.0157
	±0.005	±0.01	±0.03	±0.2	±0.2	±0.001
$\overline{B}^0$Tag	1.0997	0.665	0.840	5.12	-1.157	0.0166
	±0.006	±0.01	±0.03	±0.2	±0.2	±0.002
Lepton Tag	1.094	0.756	0.690	3.90	-1.122	0.0197
	±0.01	±0.03	±0.08	±0.3	±0.3	±0.005
Kaon Tag	1.1154	0.585	0.809	5.25	-1.533	0.0157
	±0.007	±0.02	±0.03	±0.3	±0.3	±0.002
NT1 Tag	1.118	0.786	1.036	8.9	-0.44	0.0069
	±0.01	±0.02	±0.09	±1	±1	±0.002
NT2 Tag	1.0945	0.665	0.864	6.14	-1.009	0.0152
	±0.010	±0.02	±0.05	±0.5	±0.5	±0.003
NT3 Tag	1.0648	0.654	0.868	4.95	-1.667	0.0232
	±0.007	±0.02	±0.04	±0.2	±0.2	±0.003
Perfect Tag	1.1025	0.6661	0.811	5.08	-1.560	0.0167
	±0.004	±0.009	±0.02	±0.2	±0.2	±0.001
Imperfect Tag	1.0775	0.642	0.933	6.30	-1.163	0.0152
	±0.007	±0.02	±0.03	±0.4	±0.4	±0.002
Unmixed	1.0940	0.672	0.819	5.31	-1.756	0.0148
	±0.005	±0.01	±0.03	±0.2	±0.2	±0.001
Mixed	1.1012	0.639	0.893	5.61	-1.050	0.0182
	±0.007	±0.01	±0.03	±0.3	±0.3	±0.002

Table 5.4: Results of fitting a GExp+G resolution model to $\delta\Delta t$ and $\sigma_{\Delta t}$ for signal MC.
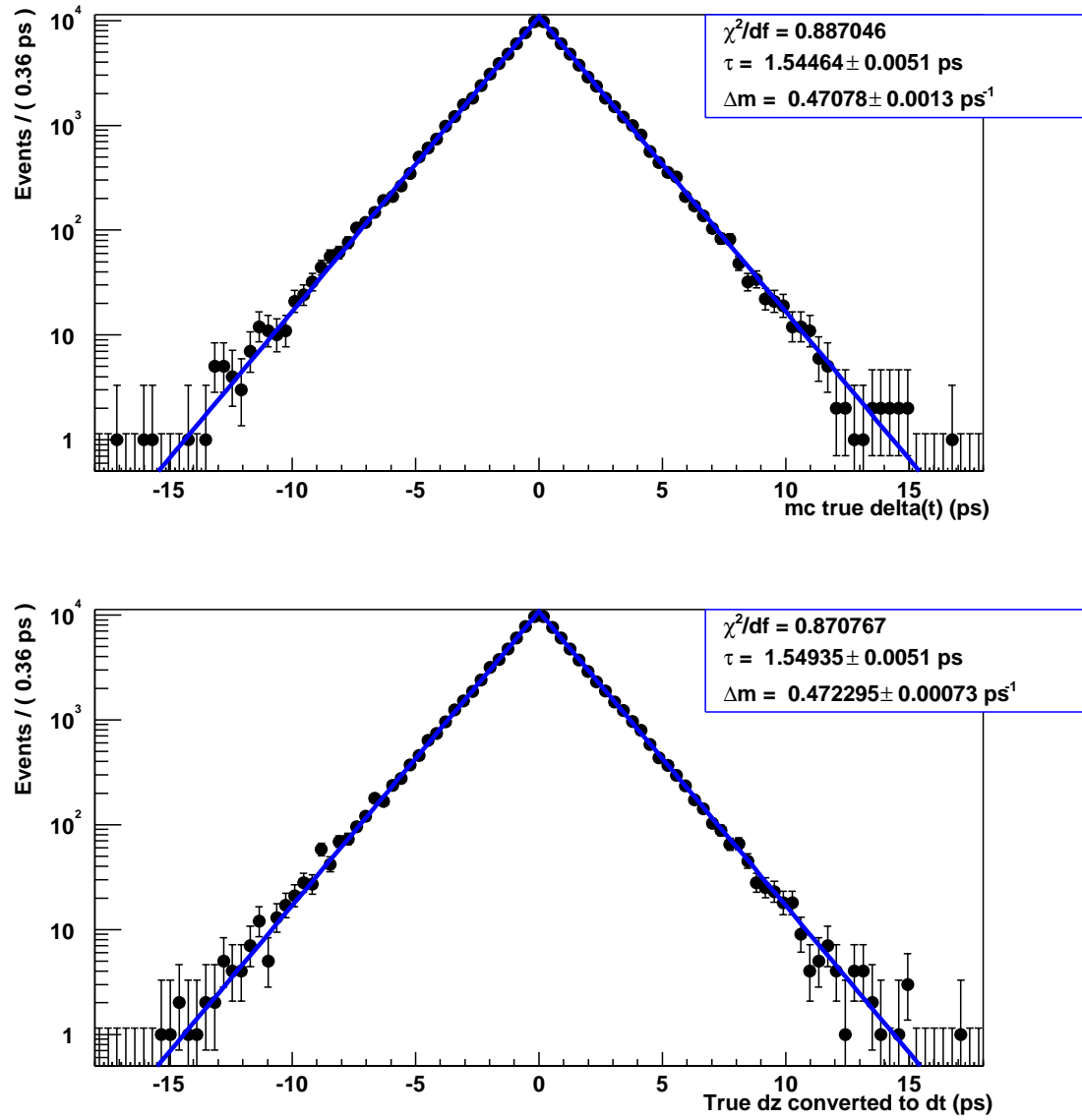
Figure 5.16: Fits to signal MC truth Δt and Δz distributions. See the text for details.
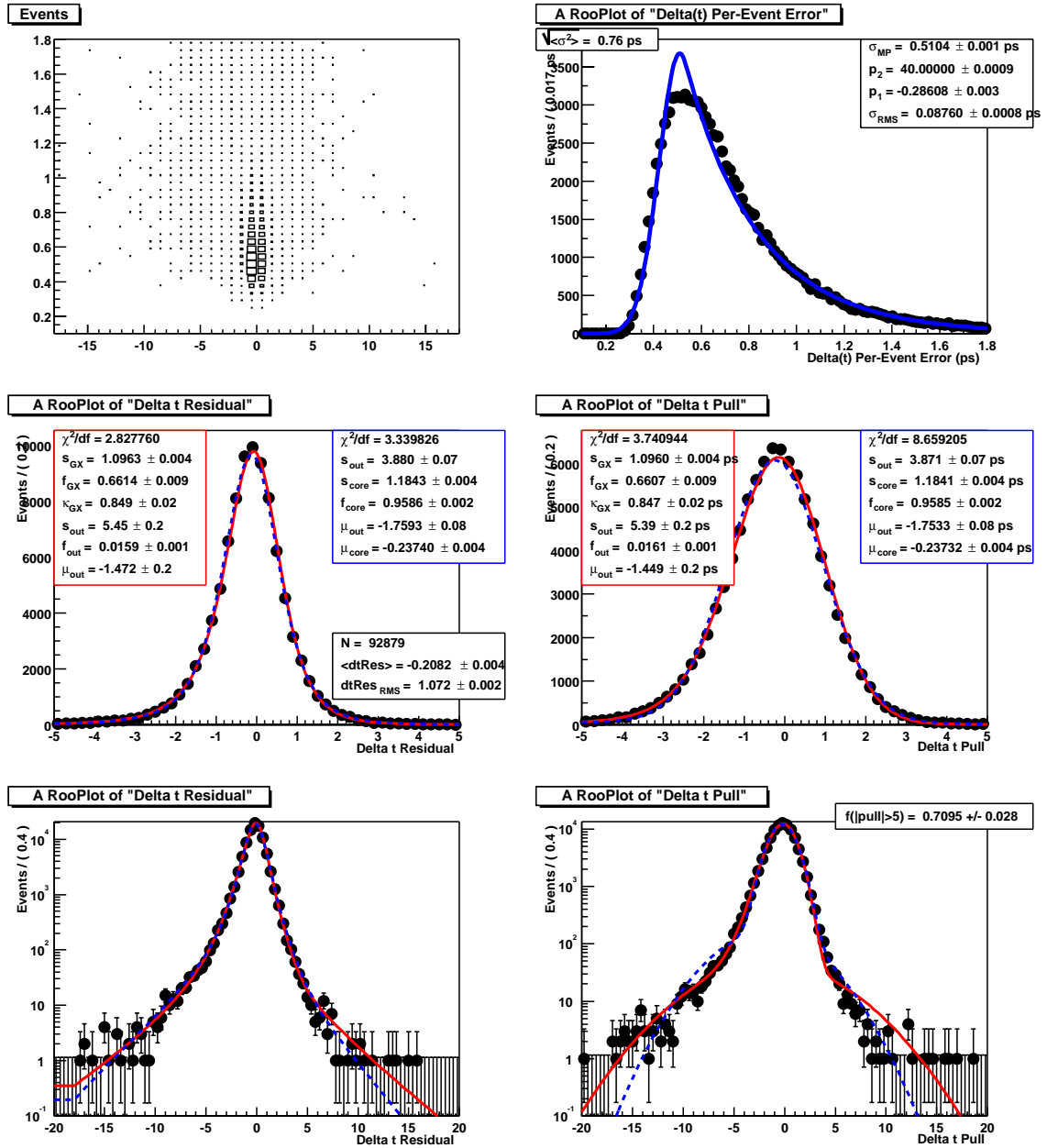
Figure 5.17: Fits to signal MC residuals and errors, for G+G model (solid curve and parameters in right-hand box) and Gexp+G model (dashed curve and parameters in left-hand box).
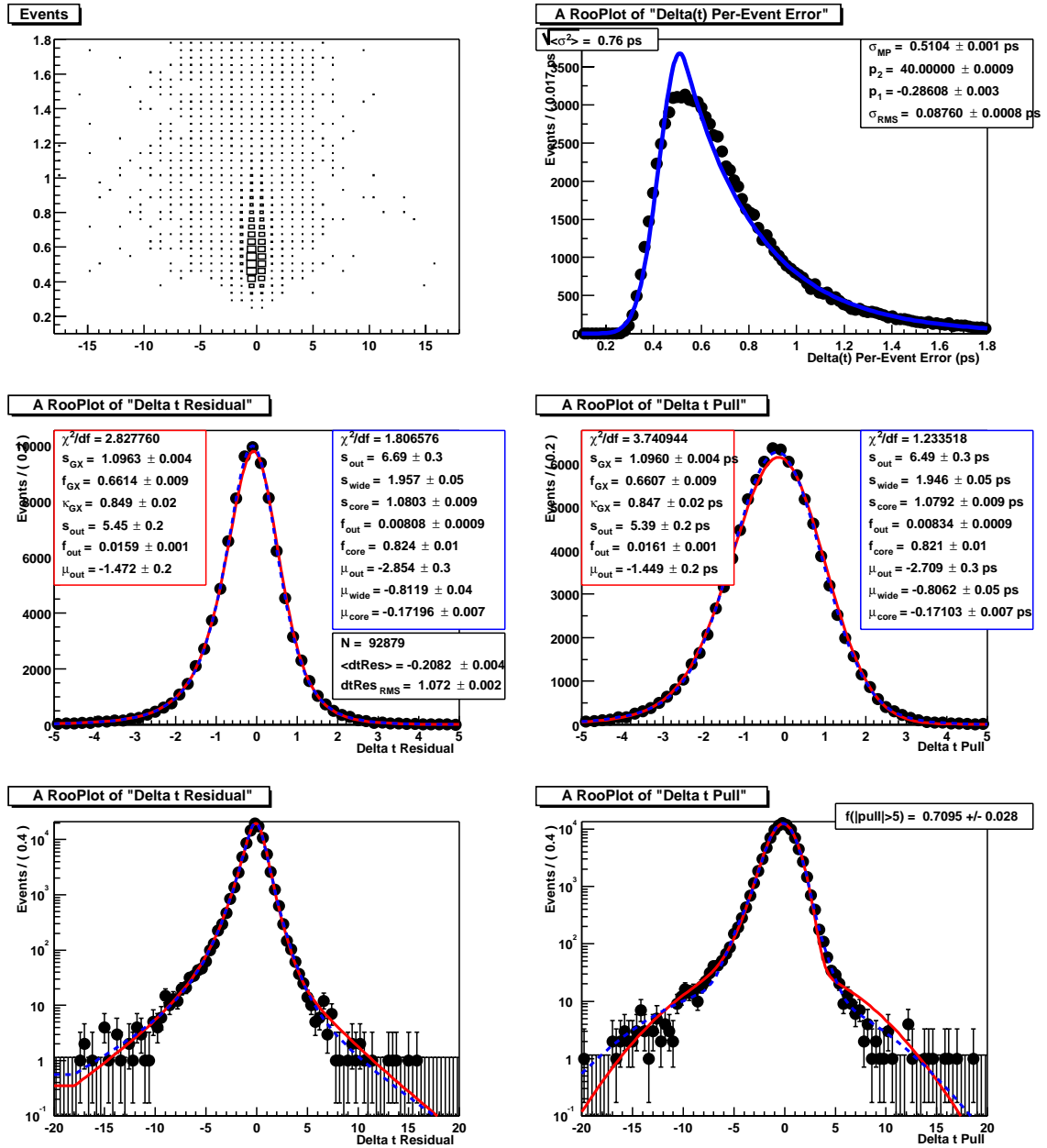
Figure 5.18: Fits to signal MC residuals and errors, for G+G+G model (solid curve and parameters in right-hand box) and Gexp+G model (dashed curve and parameters in left-hand box).
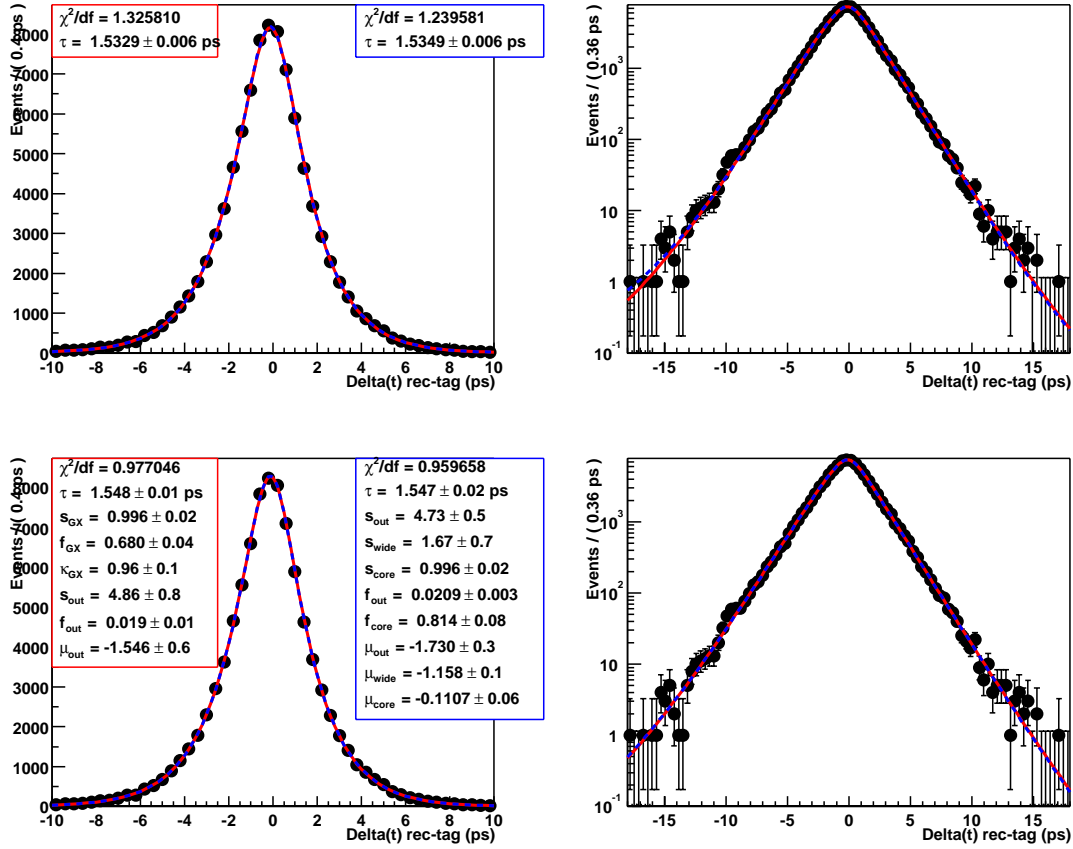
Figure 5.19: Lifetime fits to reconstructed Δt in signal MC, for G+G+G model (solid curves and parameters in right-hand box) and GExp+G model (dashed curves and parameters in left-hand box). In the top plots, the lifetime is the only free parameter; in the bottom plots, the resolution parameters are fit as well.
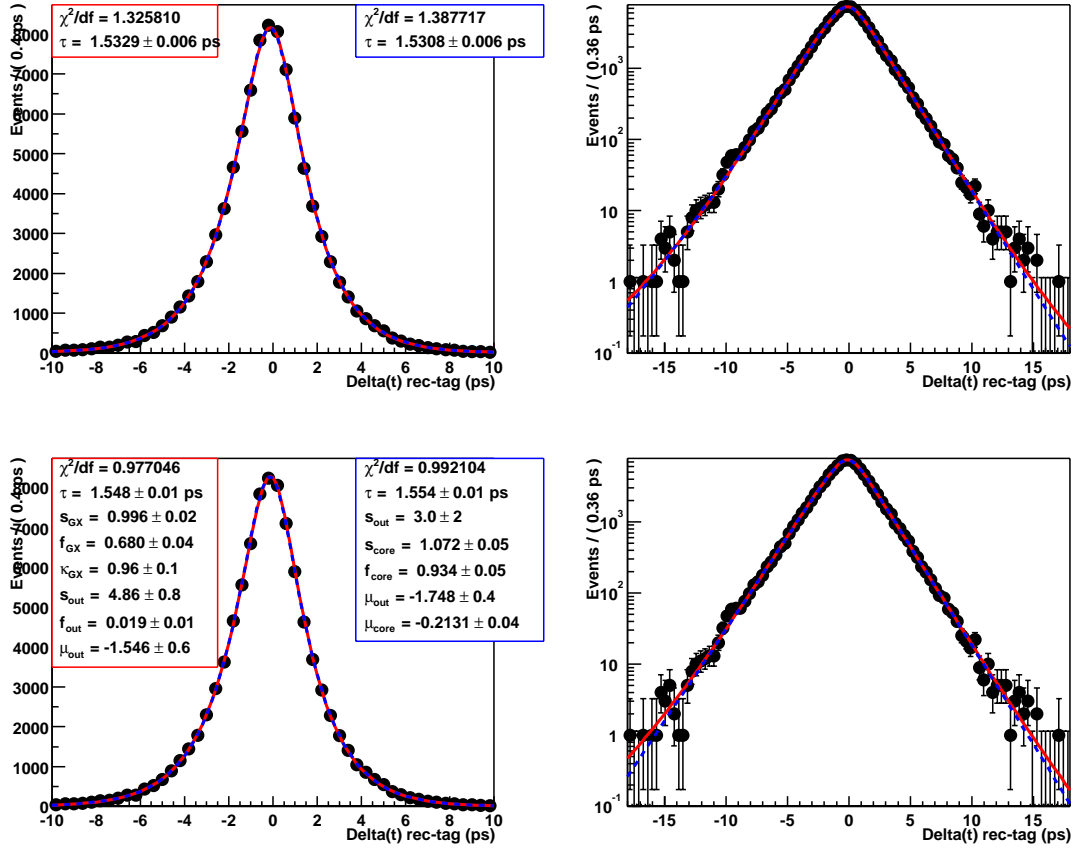
Figure 5.20: Lifetime fits to reconstructed Δt in signal MC, for G+G model (solid curves and parameters in right-hand box) and GExp+G model (dashed curves and parameters in left-hand box). In the top plots, the lifetime is the only free parameter; in the bottom plots, the resolution parameters are fit as well.

Sample	True $\Delta\tau$	True Δz	G+G+G		GExp+G	
			$\Delta\tau$ Only	$\Delta\tau$+Resln	$\Delta\tau$ Only	$\Delta\tau$+Resln
Combined	-3.361 ± 5.071	1.346 ± 5.087	-9.76 ± 5.88	2.01 ± 15.71	-11.77 ± 5.88	3.43 ± 10.19
Electrons	1.114 ± 7.502	6.457 ± 7.528	-12.18 ± 8.66	-17.97 ± 7.55	-14.21 ± 8.66	-16.9 ± 14.5
Muons	-10.8 ± 7.604	-7.502 ± 7.621	-11.96 ± 8.81	4.29 ± 17.7	-13.83 ± 8.81	7.18 ± 15.8
SVT only	-3.077 ± 5.666	0.979 ± 5.681	-10.75 ± 6.57	5.33 ± 11.52	-12.81 ± 6.57	7.41 ± 11.35
SVT+DCH	-4.517 ± 11.37	3.003 ± 11.43	-5.97 ± 13.12	-18.21 ± 16.67	-7.64 ± 13.13	-17.1 ± 24.3
$D^0 \to K\pi$	-8.39 ± 6.766	-4.15 ± 6.785	-4.93 ± 7.85	9.09 ± 12.92	-6.84 ± 7.85	11.06 ± 7.23
$D^0 \to K\pi\pi\pi$	-2.76 ± 10.52	2.244 ± 10.55	-17.81 ± 12.1	1.94 ± 31.91	-20.3 ± 12.1	-1.59 ± 48.82
$D^0 \to K\pi\pi^0$	6.099 ± 11.88	11.8 ± 11.92	-13.1 ± 13.8	0.97 ± 22.81	-15.1 ± 13.8	-1.37 ± 23.13
$D^0 \to K_S\pi\pi$	34.05 ± 33.09	38.99 ± 33.19	-16.96 ± 38.3	-40.0 ± 55.0	-18.2 ± 38.3	-43.5 ± 63.1
Reco'd B^0	-5.844 ± 7.103	-1.899 ± 7.123	-7.57 ± 8.25	-11.88 ± 15.32	-9.68 ± 8.25	-9.64 ± 12.65
Reco'd $\overline{B}^0$	-0.827 ± 7.242	4.653 ± 7.268	-11.99 ± 8.37	13.71 ± 12.4	-13.91 ± 8.37	12.76 ± 12.69
B^0Tag	-1.25 ± 7.09	4.683 ± 7.117	-11.08 ± 8.22	6.05 ± 21.68	-12.96 ± 8.22	7.14 ± 14.38
$\overline{B}^0$Tag	-5.602 ± 7.257	-2.148 ± 7.273	-8.34 ± 8.40	1.01 ± 8.24	-10.5 ± 8.40	-0.57 ± 14.32
Lepton Tag	-32.86 ± 13.69	-28.5 ± 13.72	-30.9 ± 15.2	-25.3 ± 16.08	-32.7 ± 15.2	-25.1 ± 12.6
Kaon Tag	-0.464 ± 8.588	4.969 ± 8.619	-2.44 ± 9.44	2.46 ± 4.97	-7.78 ± 9.90	5.92 ± 7.90
NT1 Tag	-11.93 ± 17.48	-7.764 ± 17.53	-10.6 ± 19.7	-15.6 ± 23.4	-12.1 ± 19.7	11.84 ± 27.04
NT2 Tag	4.918 ± 13.41	9.087 ± 13.45	-24.4 ± 15.5	-37.6 ± 27.4	-27.1 ± 15.5	-37.7 ± 27.3
NT3 Tag	4.97 ± 9.467	9.239 ± 9.493	5.37 ± 11.35	46.57 ± 22.69	3.23 ± 11.35	48.32 ± 19.01
Perfect Tag	-4.714 ± 5.902	0.743 ± 5.923	-2.44 ± 1.41	4.71 ± 9.77	-15.54 ± 6.79	-1.67 ± 11.94
Imperfect	0.411 ± 9.912	2.995 ± 9.929	1.98 ± 11.72	35.24 ± 15.89	-0.58 ± 11.72	21.56 ± 18.61

Table 5.5: Results of fitting lifetimes to various resolution models. In the first column of results, we give the difference between the lifetime extracted from a fit to the true Δt distribution and the generated lifetime (1548 fs). The second column is the difference between the lifetime extracted from the true Δz distribution converted to fs using the boost approximation and the generated lifetime. In all the other columns of results, we give the difference between the lifetime extracted from a fit to reconstructed Δt and the fit to true Δt (column 1, not the generated value of 1548 fs). For the $\Delta\tau$ Only column, the parameters for the resolution function are fixed to the values extracted from the fit to residuals and errors; for the $\Delta\tau$+Resln column the resolution parameters are floating in the fit. All results are in fs.
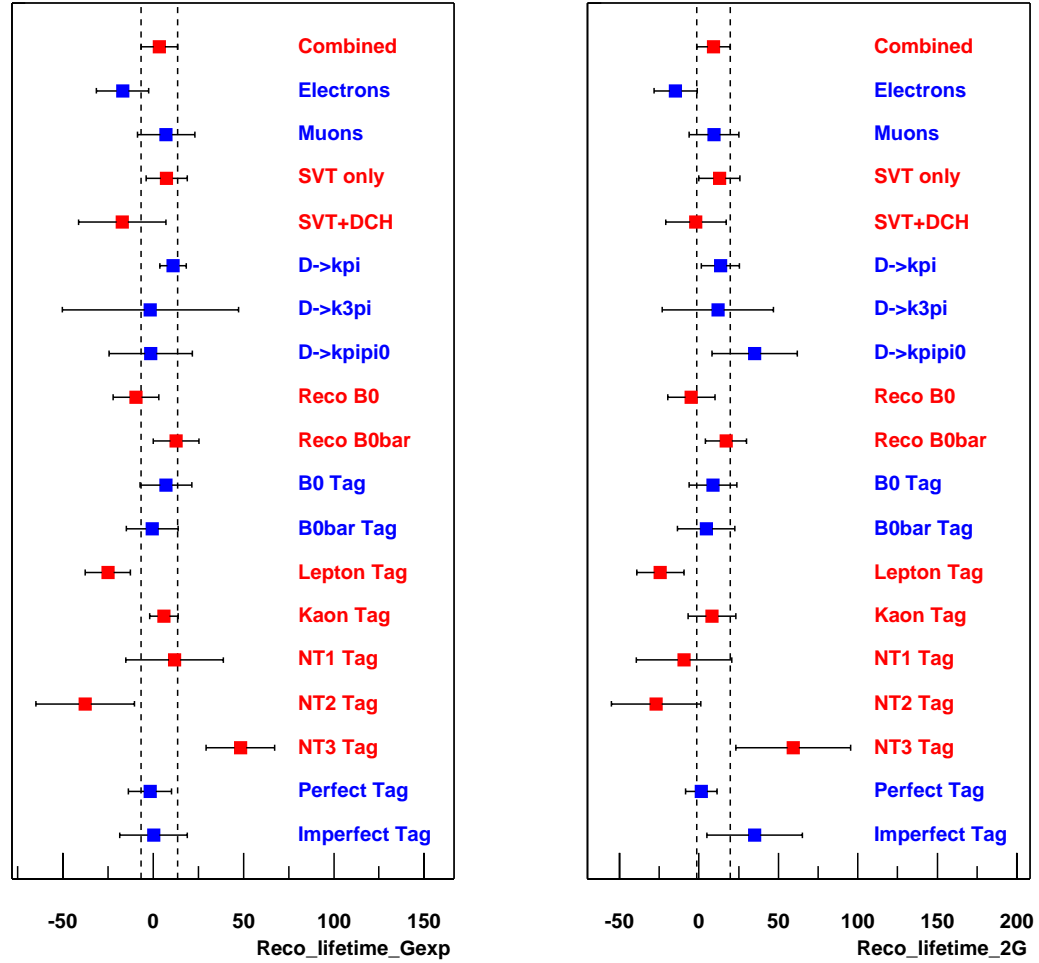
Figure 5.21: Comparison of lifetimes fitted with the GExp+G (left-hand side) or G+G (right-hand side) resolution models. Both the lifetime and the resolution model parameters are free in the fit. The displayed values are (fitted value − MC truth) in ps.
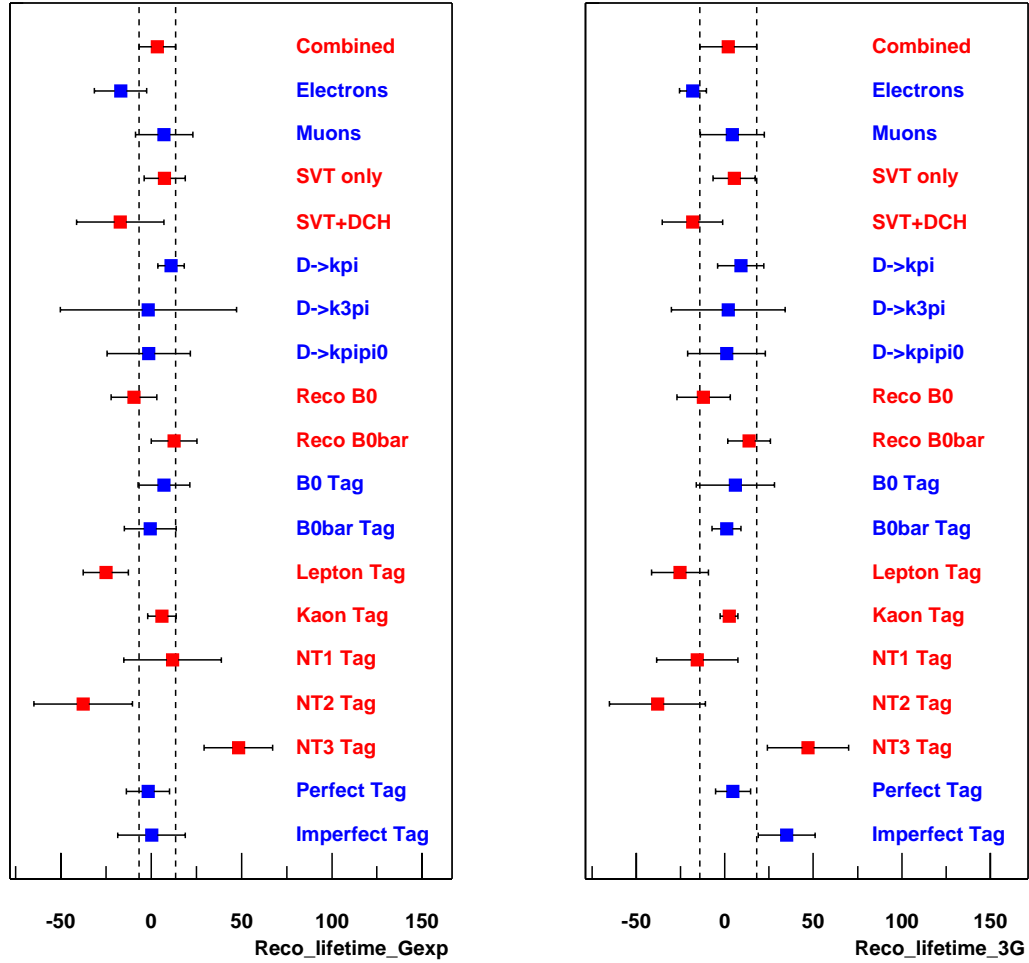
Figure 5.22: Comparison of lifetimes fitted with the GExp+G (left-hand side) or G+G+G (right-hand side) resolution models. Both the lifetime and the resolution model parameters are free in the fit. The displayed values are (fitted value − MC truth) in ps.

Chapter 6

Flavor Tagging

6.1 Introduction to tagging

Tagging the opposite side B^0 flavor is the third critical step in the measuring Δm_d. The flavor of the reconstruction side is easily determined from the charge of the lepton/D^*. Therefore, to determine the mixing status of the event, one must determine the flavor of the opposite side B^0.

BaBar employs two very different but complementary strategies for tagging the flavor of the opposite side B^0, one is a cut based algorithm (NOT - Non Optimal Tagging) while the second is a neural network based approach (NetTagger). The ultimate approach used for this thesis and by many analyses in BaBar is a hybrid which retains the best qualities of both methods.

6.2 Cut Based Tagging

NOT is a cut based tagging algorithm. NOT uses identified electrons, muons, kaons and slow pions to determine the opposite side's flavor. The algorithm receives

all charged tracks not used to reconstruct the $D^*l\nu$ candidate and uses PID discussed in section 4.5 to determine the particle types on the tag side of the event. NOT then tries to use the following correlations between particle types and B^0 flavor [20].

1. Lepton: A high momentum lepton ($p^* >$1.1 GeV) typically comes directly from the B^0, thus the lepton charge is strongly correlated with the B^0 flavor and is considered to be the best tag available. Medium momentum leptons (0.6$< p^* <$1.1 GeV) also carry useful information, but they require some additional quality cuts (depending on lepton type) and suffer from a larger mistag rate than high momentum leptons. At low momentum ($0.3 < p^* < 0.5$ GeV) cascade leptons dominate and the reverse sign can be used for tagging. This only works for electrons since there is no muon id at this low a momentum. See Figure 6.1 for a depiction of primary and secondary leptons.

2. Kaon: The net charge of identified kaons is correlated to the B^0 flavor through the most probable b quark decay chain (b→c→s). This correlation isn't a strong as the high momentum lepton correlation due to "wrong sign decays", but it is still very good and is extremely important due to the large numbers of events which contain kaons on the tag side of the event.

3. slow pion: Slow pions come from the decay $D^* \rightarrow D^0\pi^+$. The slow pion determines the flavor of the D^* which is correlated to the flavor of the B^0. Correctly identifying slow pions can be difficult due to large backgrounds, thus slow pions are not as useful as leptons or kaons.

4. p^*:If there are no other tags available, the charge of the highest momentum particle can be used, assuming $p^* > 1.1$ GeV. This uses the idea that you may have found a

lepton which didn't pass the lepton PID cuts, but still carries tagging information. This tag is very weak and is used only when no other information is present.
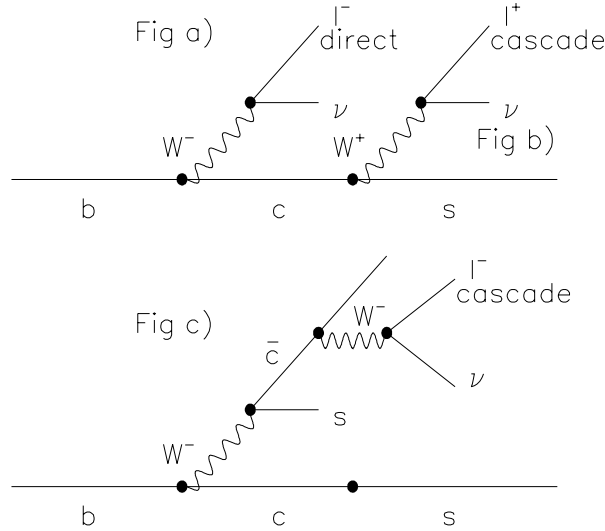
Figure 6.1: Figure a shows a direct lepton while b and c show cascade leptons. The direct leptons tend to have a much harder momentum spectrum [26].

Quite often there will be more than one piece of tagging information available. In such cases, the tagging information could conflict. If a lepton tag and kaon tag are both present but conflict, the event will not be tagged. Since the lepton and kaon tags are much more reliable than the slow pion or p^* tags, any conflict with the two low priority tags will be ignored. There is no reason to discard a good lepton tag because a suspect slow pion tag conflicts. All NOT tagged events will fall into one of seven different categories (list in decreasing order of priority) [20]:

- **ElKaon**: There is both an electron tag and kaon tag and they agree. If they disagree the event will not be tagged by NOT.

- **MuKaon**: There is both an muon tag and kaon tag and they agree. If they disagree the event will not be tagged by NOT.

- **Electron**: There is an electron tag present, no kaon tag.

- **Muon**: There is a muon tag present, no kaon tag.

- **Kaon**: There is a kaon tag present, no lepton tag.

- **SlowPion**: If there is a slow pion satisfying certain cuts, use it to tag the B^0.

- **NoPid**: If there is a maximum p^* tag, use it to tag the B^0.

Thus an event will not be tagged if there are no tags at all or a lepton/kaon tag conflicts.

One of the inherent strengths of NOT is that each category is based on an easy to understand physics process.

6.3   NetTagger

NetTagger is a neural network based tagging algorithm. Each event is put into a series of neural networks, each one looks to test the consistency of the event with a particular method of tagging. Each of these networks is called a Feature Net. The Feature Nets use a number of variables taken from the event including electron, muon and kaon PID, track momentum, calculated energies and angles involving recoil W's, track POCA and z0 positions, etc... NetTagger currently uses four Feature Nets:

- <u>L-net</u>: This network determines how compatible a track is with an event where the track came directly from a b quark semi-leptonic decay. This net uses electron and muon PID, plus various kinematic quantities related to the virtual W producing the lepton. This net essentially looks for primary leptons.

- <u>LS-net</u>: This network determines how compatible a track is with being a lepton, but not the primary lepton.

- <u>K-net</u>: This network determines how compatible a track is with being a kaon using track momentums plus kaon and pion likelihoods from the SVT, DCH and DIRC.

- <u>Soft-pion-net</u>: This network determines how compatible a track is with being a soft-pion.

Once the Feature Nets have been created, they are all combined into a single output probability (ranging from 0 to 1) describing whether an event is like a $\overline{B}^0$ (NetTagger = 0) or a B^0 (NetTagger = 1). Events with a value of 0.5 are undetermined; NetTagger has no preference for calling such an event a B^0 or $\overline{B}^0$. There are no categories or any indication as to which Feature Net was used, just a single output probability.

The strength of NetTagger is that every little bit of information is used to squeeze out extra performance. NetTagger has always had a slightly better tagging performance when measured against cut based tagging methods.

6.4 Elba Tagger

The decision of which tagging algorithm to use is very difficult. The tagging tools developers have made very detailed comparisons between the performance of Net-Tagger and NOT to discover where NetTagger gets its extra tagging performance. The ultimate conclusion is that NOT and NetTagger have very comparable performance for kaon and lepton tags. These tagging categories are relatively straight forward and don't significantly benefit from extra correlation information available in NetTagger. NetTagger clearly out-performs NOT in areas involving slow pions, unidentified high momen-

tum tracks and untagged events (for example conflicting lepton/kaon tags). This realization lead to a hybrid called Elba Tagger.

The Elba Tagger uses NOT categories where the tagging performance between NOT and NetTagger is essentially equal to retain the clear physics understanding of the underlying events and uses NetTagger for the remaining events to optimize the tagging performance. The events tagged by NetTagger are broken into two different categories, Net1 and Net2. This division groups together events that are relatively close in mistag rate. It is important that events within a single category be somewhat uniform in terms of Δt resolution and mistag rate since a single average value for these quantities will be extracted during the final fitting process. If there are significant variations, then the final extracted values for resolution parameters or mistag rates could suffer from biases created by the poor approximation.

	Bins for NetTagger probability	
Category	for $\overline{B}^0$ tag	for B^0 tag
NetTagger 1 (Net1)	0.0 - 0.25	0.75 - 1.0
NetTagger 2 (Net2)	0.25 - 0.4	0.6 - 0.75

Table 6.1: NetTagger categories used by the Elba Tagger.

The Elba Tagger contains the following categories:

- **Lepton**: This category is the sum of the NOT categories ElKaon, MuKaon, Electron and Muon. These categories all have very similar mistag rates and underlying physics (semi-leptonic decay of the b quark). Breaking this category down into smaller categories is possible, but results in no significant improvement. This category is the best one in terms of mistag rate.

- **Kaon**: This is the NOT Kaon category, this category is the largest category and

provides the bulk our Elba Taggers tagging power.

- **Net1**: All events not falling into the Lepton or Kaon categories and having a Net-Tagger output in the range shown in table 6.1

- **Net2**: All events not falling into the Lepton or Kaon categories and having a Net-Tagger output in the range shown in table 6.1

Figure 6.2 shows the relative number of events falling into the two different NetTagger categories. A third category could be created for events between 0.4 and 0.6, but this would add very little to the tagging power of Elba Tagger, these events suffer from very high mistag rates which very quickly dilutes their effectiveness.
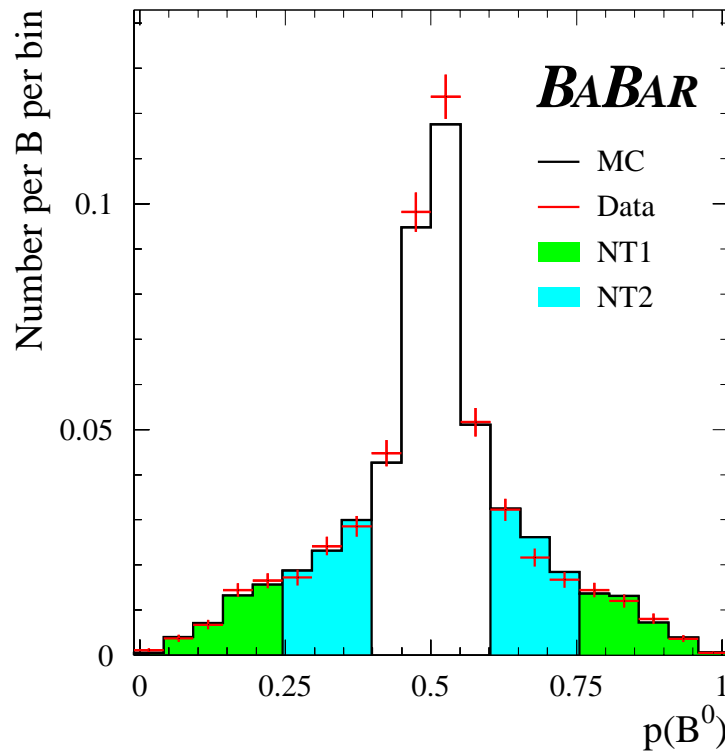
Figure 6.2: NetTagger output broken into categories. The white zone represents events which are not tagged by Elba Tagger [21].

6.5 Tagging in this Analysis

This analysis uses a custom made tagger which adds a fifth category to the Elba Tagger, Net3. Net3 is composed of all the events not used by the previous four categories. These events have extremely little tagging power, the real benefit is derived from the increased number of events used to determine resolution parameters.

6.6 Tagging performance

Tagging performance is important to define quantitatively so that it can easily be measured and used to determine the best approach for tagging. The figure of merit used for tagging is known as the total separation, Q:

$$Q_{tot} = \sum_i Q_i = \sum_i \epsilon_i D_i^2 = \sum_i \epsilon_i (1 - 2\omega_i)^2 \tag{6.1}$$

where ω_i is the mistag rate for the ith tagging category, ϵ_i is the tagging efficiency for the ith tagging category (i.e. N_i/N_{total}) and D_i = 1-$2\omega_i$ is known as the dilution for the ith tagging category. The dilution takes into account the fact that getting a tag wrong counts against you twice. If you only get half the events right, you have no tagging power: you could get the same results with random guessing. The total separation has a close connection to the expected statistical error on the Δm_d measurement.

The probability distribution function which describes the time dependence of the $B^0\overline{B}^0$ oscillations from section 1.4 is given by:

$$f_i = \frac{1}{4\tau} e^{-|\Delta t|/\tau} (1 + gD_i \cos(\Delta m \Delta t)) \tag{6.2}$$

where g is +1 for unmixed events, -1 for mixed events and 0 for untagged events. The

error associated with Δm_d can be determined using the well known relation [10]:

$$\sigma^{-2} = N \int dx \frac{1}{f} \left(\frac{\partial f}{\partial \alpha} \right)^2 \tag{6.3}$$

where f(x;α) is normalized as

$$\int dx f(x; \alpha) = 1 \tag{6.4}$$
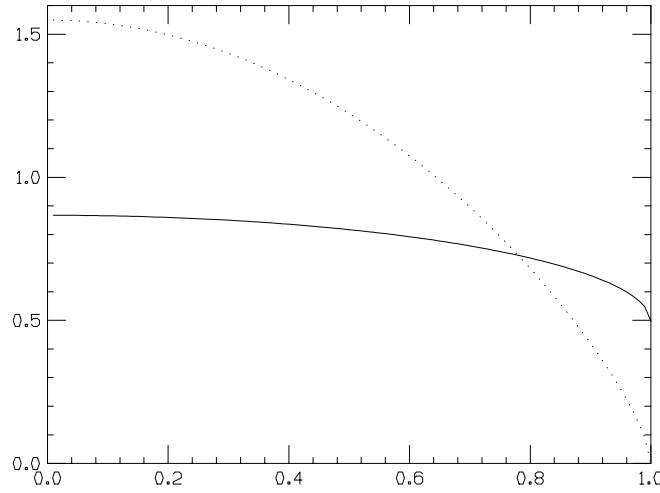
Figure 6.3: The solid curve is F($\Delta m\tau$,D) evaluated numerically. The horizontal axis is dilution, $\Delta m\tau$ is fixed at 0.71. [10]

for a single tagging category, the error on Δm_d is [10]:

$$\sigma(\Delta m) = F(\Delta m\tau, D_i)^{1/2} \frac{1}{\tau D_i \sqrt{N_i}} = F(\Delta m\tau, D_i)^{1/2} \frac{1}{\tau D_i \sqrt{\epsilon_i} \sqrt{N}} \tag{6.5}$$

F($\Delta m\tau$,D) is 1 to first order, with correction terms with depend on D and $\Delta m\tau$. A plot of F

is shown in figure 6.3 with the value of $\Delta m\tau$ fixed at 0.71 (PDG 2000 value is 0.73±0.029).

The possible variations due to the $\Delta m\tau$ term are significantly smaller than the known

variations due to the dilution. The dilution ranges from about 0.85 (Lepton category)

down to about 0.3 (Net2). Taking the square root of this plot will reveal an even flatter

curve closer to 1. When resolution effects are taken into account, the overall picture

doesn't change (but there are no longer analytic solutions to several of the intermediate

steps). The final result is that the total error on Δm_d using all tagging categories goes as

$$\sigma(\Delta m) \approx \frac{const}{\sqrt{Q_{tot} N}} \tag{6.6}$$

This leads to the simple interpretation that the N tagged events used for the Δm_d measurement has the same statistical power of Q_{tot}N perfectly tagged events.

Table 6.2 shows the value of Q_{tot} obtained from signal Monte Carlo. The tagging efficiency per category is obtained by using a simple counting of events in category i divided by the total number of events. The mistag rate is measured simultaneously with the B^0 lifetime and mixing frequency.

Category	ϵ_i	ω_i	D_i	$Q_i = \epsilon_i D_i^2$
Lepton	13.2	.063	.874	10.1
Kaon	35.1	.157	.686	16.5
Net1	8.3	.193	.614	3.1
Net2	14.5	.338	.324	1.5
Net3	28.9	.465	.070	0.14
Total	100			29.9

Table 6.2: Elba tagger + NT3 performance on signal Monte Carlo $D^* l \nu$ data.

6.6.1 B^0 vs $\overline{B}^0$

There is a difference between the tagging efficiency of B^0 events vs. $\overline{B}^0$ events due to the different interactions of the daughter particles and matter. If the tag side decays as a B^0, there will most likely be a K^+ daughter in the final list of daughter particles, and a K^- if the particles decays as a $\overline{B}^0$. The K^+ and K^- have very different cross sections when traveling through matter which will have a significant impact on how easily the daughters and detected and properly identified. The effect seems to be more pronounced in the NetTagger categories. The NetTagger categories are more likely

to have events that are missing particles or fail to have a clearly identified Kaon present. The lower the NetTagger precedence, the more pronounced the effect. Thus NT3 has the largest measured mistag rate difference of any category.

6.7   Correlation between Mistag rate and $\sigma_{\Delta t}$

A surprising and unexpected correlation has been discovered between mistag rate (ω) and per-event-error ($\sigma_{\Delta t}$) in the Kaon tagging category. This section explores the possible reasons for this strong correlation (shown in figure 6.4).

6.7.1   Kaon Flight Correlation

One possibility is that ω_{Kaon} is correlated to $\sigma_{\Delta t}$ through Kaon PID. It has already been shown that $\sigma_{\Delta t}$ is correlated with the charm flight polar angle (in section 5.7.2), if the kaon PID efficiency depended on the Kaon flight polar angle (which is correlated to the charm flight polar angle) there is a connection between Kaon PID and $\sigma_{\Delta t}$.

Figure 6.5 shows that Kaon PID doesn't explain the ω_{Kaon}-$\sigma_{\Delta t}$ correlation, since the slope of the line in figure 6.5 is almost unaltered when using perfect Kaon PID.

6.7.2 D^0, D^+ and D_s

A second possibility is that different charm species have different $\sigma_{\Delta t}$ distributions due to lifetime, multiplicity differences and also have different ω_{Kaon}. Figure 6.6 shows that different charm species definitely have different mistag rates, but the relative fraction of each species doesn't change much. Given the strong correlation observed between ω_{Kaon} and $\sigma_{\Delta t}$, this isn't the source.
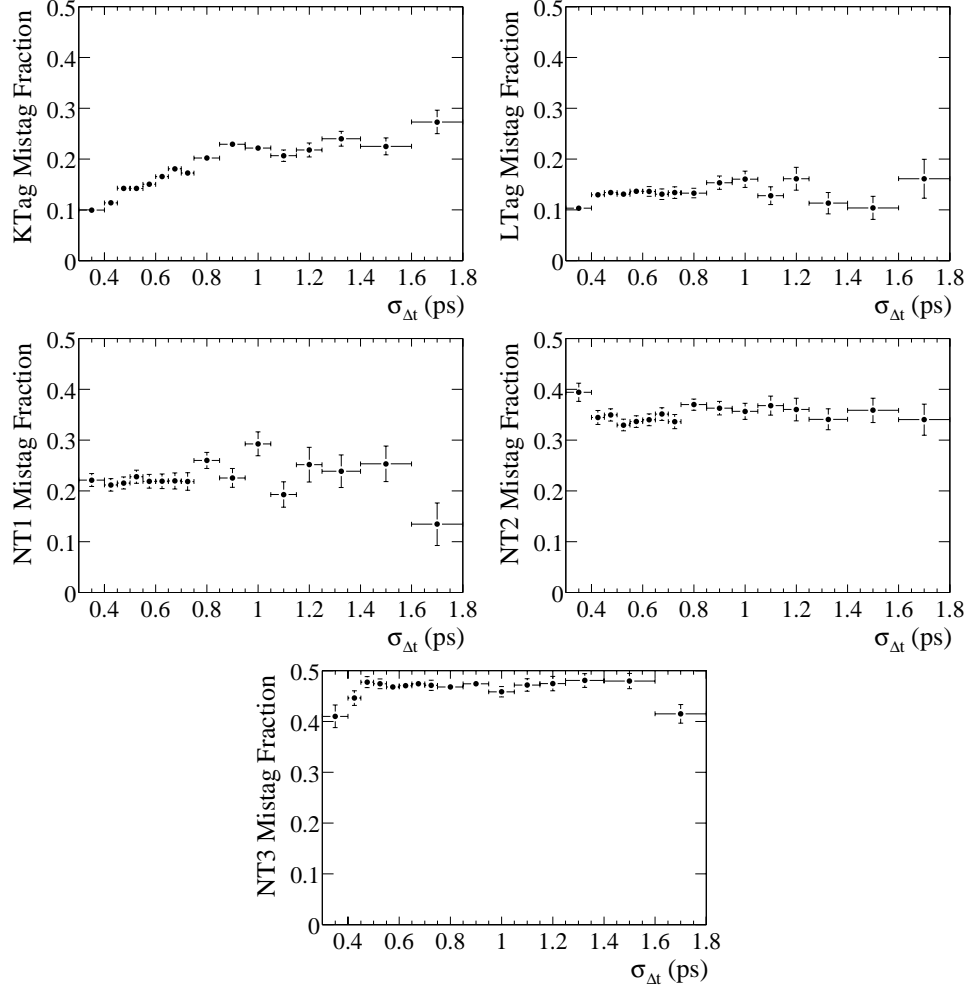
Figure 6.4: Mistag rates by tagging category.

6.7.3 Transverse Momentum

The final source considered is correlations due the transverse momenta of used charged tracks. This idea has been shown by Art Snyder [29] to be almost solely responsible for the $\omega\text{-}\sigma_{\Delta t}$ correlation. The basic idea is that a single track contributes to the z position error roughly like:

$$\sigma_z^i \propto \frac{1}{p^i \sin\theta^i} = \frac{1}{p_t^i} \qquad (6.7)$$
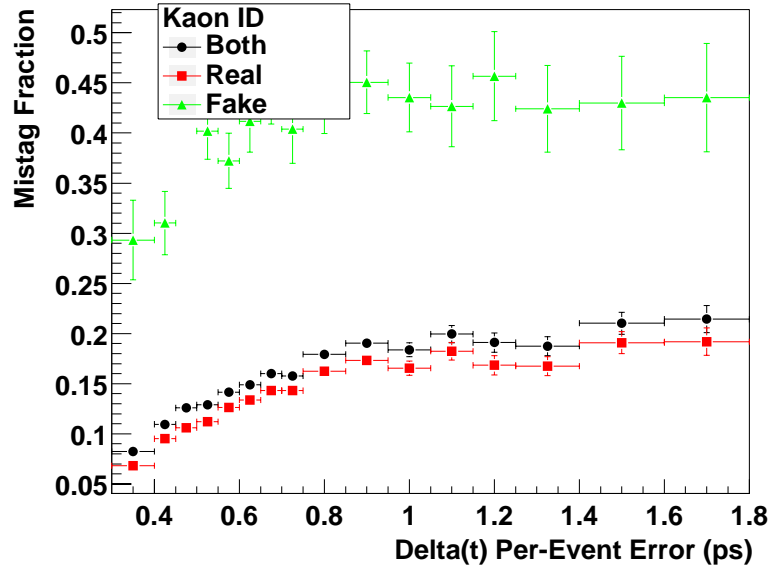
Figure 6.5: ω-$\sigma_{\Delta t}$ correlation shown for all events, events with correctly identified kaons and incorrectly identified kaons.

The total error is given by:

$$\frac{1}{(\sigma_z)^2} = \sum \frac{1}{(\sigma_z^i)^2} = \sum (p_t^i)^2 \tag{6.8}$$

Since the Δz error is dominated by the tag side tracks, $\sigma_{\Delta t}$ should be a linear function of $1/\sqrt{\sum(p_t^i)^2}$. The left side of figure 6.7 shows this relationship while the right side shows the scaling between ω and $1/\sqrt{\sum(p_t^i)^2}$. The correlation between $\sigma_{\Delta t}$ and ω_{Kaon} could
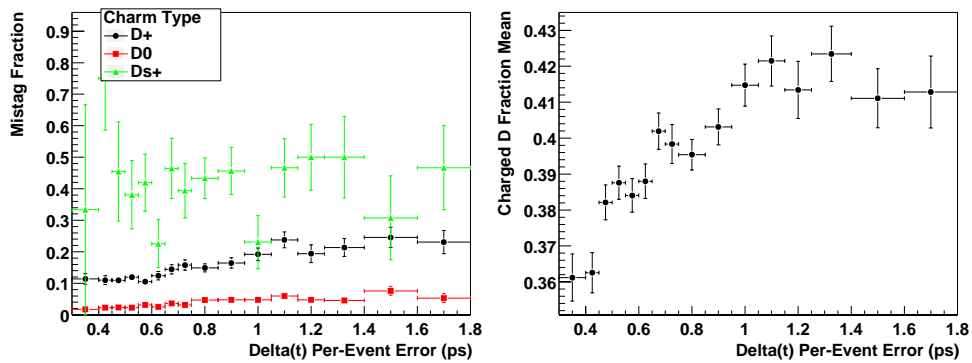
Figure 6.6: Left: ω-$\sigma_{\Delta t}$ correlation shown for D^0, D^+ and D_s events. Right: ratio of D^+ to the sum of D^0 and D^+ events.

have several different sources. In order to declare transverse momentum as the primary reason for the correlation, the correlation should mostly disappear once the transverse momentum effect is taken into account. Figure 6.8 shows the mistag correlation is almost completely removed when correcting for the p_t contribution determined from figure 6.7.

The mechanism for the correlation can be found in the $\sqrt{\sum (p_t^i)^2}$ spectrum for right and wrong sign tags shown in figure 6.8. The wrong sign tags have a softer spectrum which correlates with a larger mistag rate. The source of the softer p_t spectrum is connected with the charged track multiplicity of right and wrong sign decays. The correct decay for a D^+ is to a K^-, this requires a minimum of 3 charged tracks to conserve charge. A wrong sign version of this would be $D^+ \to K^+$ which has $K^+ \overline{K^0}^* (\pi^0)$ as the dominant decay modes. The neutral particles aren't counted in the p_t spectrum, thus these modes have a softer spectrum. The effect for D^0's is quite small, wrong sign decay modes are Cabibbo or Doubly Cabibbo suppressed. This mechanism not only accounts for the correlation, it appears to dominate any other possible effects correlating ω_{Kaon} and $\sigma_{\Delta t}$.
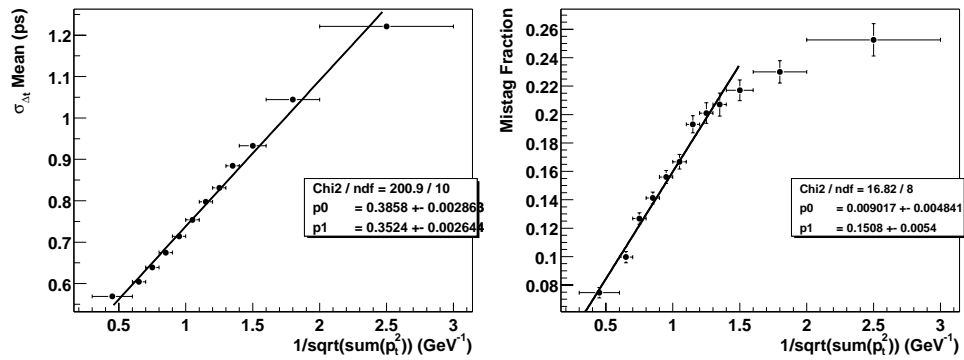
Figure 6.7: Mean value of $\sigma_{\Delta t}$ (left) and ω (right) as a function of $1/\sqrt{\sum (p_t^i)^2}$.
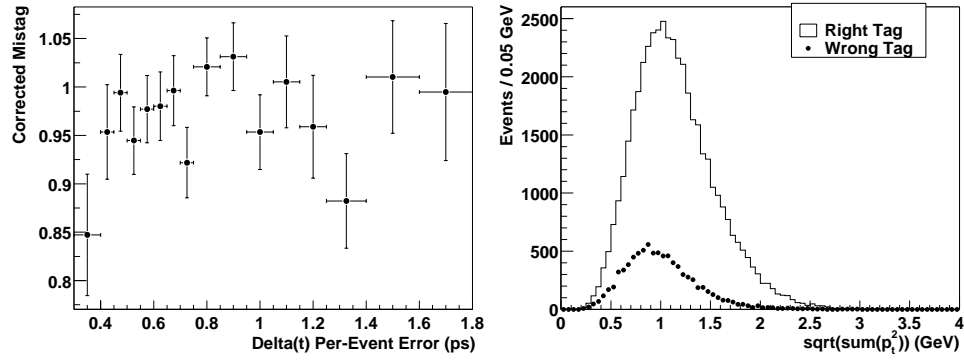
Figure 6.8: Left: Corrected mistag, ω/ω_{calc}, where $\omega_{calc} = 0.151/\sqrt{\sum(p_t^i)^2} + 0.009$. Right: $\sqrt{\sum(p_t^i)^2}$ spectrum for right and wrong tagged events.

Chapter 7

Fit Method

This analysis uses an unbinned maximum likelihood approach to fit for the B^0 lifetime, mixing frequency and mistag rates. This approach entails modeling the data set with one or more probability distribution functions (PDF's) which describe the underlying distributions. The PDF's are a function of one of more unknown variables that are varied to determine what value of each particular variable best describes the complete data set, given the PDF's that have been chosen. This is accomplished by forming the likelihood function:

$$\mathcal{L} = \prod P(\vec{x}; \vec{a}) \tag{7.1}$$

where P is the PDF chosen, $\vec{x}$ represents measured quantities in the data set (for example, D^* - D^0 mass, Δt and $\sigma_{\Delta t}$) and $\vec{a}$ represents the unknown variables we want to measure (Δm_d, τ and ω for example).

The PDF used in forming the likelihood function is generally a combination of several PDF. For this analysis, the general strategy is to create a PDF of the form:

$$P(\delta m, \Delta t, \sigma_{\Delta t}, g; \vec{a}) \quad = \quad f^{back} \cdot F^{back}(\delta m; \vec{p}^{back}) \cdot G^{back}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{back})$$

$$+ (1 - f^{back}) \cdot F^{sig}(\delta m; \vec{p}^{\,sig}) \cdot G^{sig}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,sig}) \quad (7.2)$$

where g is the mixing status of the event, δm represents the D^* - D^0 mass difference, F is a PDF which describes the probability of an event being signal or background and G is a PDF which describes the Δt structure of signal and background events.

This is a quick sketch of the final PDF used for this analysis, subsequent sections will fully describe the PDF's F and G and will introduce categories used to further improve the ability of the final PDF to describe the data sample.

7.1   Data Categories

The $D^* l\nu$ dataset is broken down into different categories to either define new control samples (on/off resonance defines the off-peak control sample for example) or because a physics parameter varies with respect to a category, for example mistag rate and tagging category. Table 7.1 lists the different categories used for this analysis.

The first three categories are used to define different control samples for backgrounds. The AngCut is used to define if an event is in the signal sample (Opposite side) or flip lepton sample (Same Side). OnOffRes reveals whether the event was taken on the Υ(4s) resonance or off-peak. The off-peak sample is a control sample for $c\bar{c}$ continuum events. Lepton ID determines if the event is in the signal sample (good electron or muon) or part of the fake lepton control sample.

The next three categories (SVTDCH, D0mode, TagCat) are defined because various fit parameters change significantly across sub-categories. This implies that using sub-categories can give a significant gain in statistical power (smaller statistical errors in a fit) over just using an average value over all sub-categories.

Category	N	Description
AngCut	2	What is the value of $\cos\theta_{D^*l}$, $\cos\theta_{B,D^*l}$?
		(1) Opposite-side (OS) events have $\cos\theta_{D^*l} <0$, $|\cos\theta_{B,D^*l}| <1.1$
		(2) Same-side (SS) events have $\cos\theta_{D^*l} >0$, $|\cos\theta_{B,D^*-l}| <1.1$
OnOffRes	2	What was the center of mass energy for the event?
		(1) near the $\Upsilon(4s)$ resonance (on-peak)
		(2) well below the $\Upsilon(4s)$ resonance (off-peak)
LeptonID	3	What was the value of lepton PID?
		(1) candidate passes very tight electron PID
		(2) candidate passes very tight muon PID
		(3) candidate fails both loose electron and muon PID
SVTDCH	2	How was the slow pion candidate track reconstructed?
		(1) in the SVT only
		(2) in the SVT and DCH
D0mode	3	What decay mode was the D^0 candidate reconstructed in?
		(1) $D^0 \to K\pi$
		(2) $D^0 \to K\pi\pi^0$
		(3) $D^0 \to K\pi\pi\pi$ or $D^0 \to K_S\pi\pi$
TagCat	5	What was the tagging category used for tagging the event?
		(1) Lepton
		(2) Kaon
		(3) Net1
		(4) Net2
		(5) Net3

Table 7.1: Summary of categories used in this analysis [25].

7.2 Signal probability PDF, $F(\delta m; \vec{p})$

The signal probability function depends on the value of δm and the categories which affect the signal and combinatoric background shapes. There are three components to determining the signal probability function:

1. Fit the signal shape of correctly reconstructed D^*'s

2. Fit for the combinatoric background shape of misreconstructed D^*'s

3. Determine the peak and combinatoric yields for each sample

Sample	AngCut	OnOffRes	Lepton ID	Sample enriched in...
1	OS	On	e	electron signal
2	OS	On	μ	muon signal
3	OS	On	!e,!μ	fake BG to samples 1 and 2
4	OS	Off	e	continuum BG to sample 1
5	OS	Off	μ	continuum BG to sample 2
6	SS	On	e	uncorrelated BG to sample 1
7	SS	On	μ	uncorrelated BG to sample 2
8	OS	Off	!e,!μ	fake BG to samples 4 and 5
9	SS	On	!e,!μ	fake BG to samples 6 and 7
10	SS	Off	e	continuum BG to sample 6
11	SS	Off	μ	continuum BG to sample 7
12	SS	Off	!e,!μ	fake BG to samples 10 and 11

Table 7.2: Summary of the signal and control samples selected for lifetime and mixing analysis. Samples 1–2 are signal, 3–7 are primary background control samples, and 8–12 are secondary background control samples.

Correctly reconstructed D^* events are modeled with a double Gaussian, this is the simplest signal model which gives a good fit:

$$F_{peak}(\delta m) = \sum_{k=1}^{2} \frac{f_k}{\sqrt{2\pi}\sigma_k} exp\left(-\frac{(\delta m - \delta m_k)^2}{2\sigma_k^2}\right) \tag{7.3}$$

where $f_1 + f_2 = 1$. The combinatoric background is modeled with the function:

$$F_{comb}(\delta m) = \frac{1}{N}\left[1 - exp\left(-\frac{\delta m - \delta m_0}{c_1}\right)\right]\left(\frac{\delta m}{\delta m_0}\right)^{c_2} \tag{7.4}$$

where N is a normalization constant and δm_0 is the kinematic threshold equal to the π^+ mass. Figure 7.1 shows the shapes the background function can assume.

The background shape is forced to zero at the kinematic limit of the π^+ mass. The exponent part enforces the turn on at the π^+ mass and asymptotically approaches the value 1 for large values of δm. The power law portion allows the PDF to match the long combinatoric tail at large values of δm.
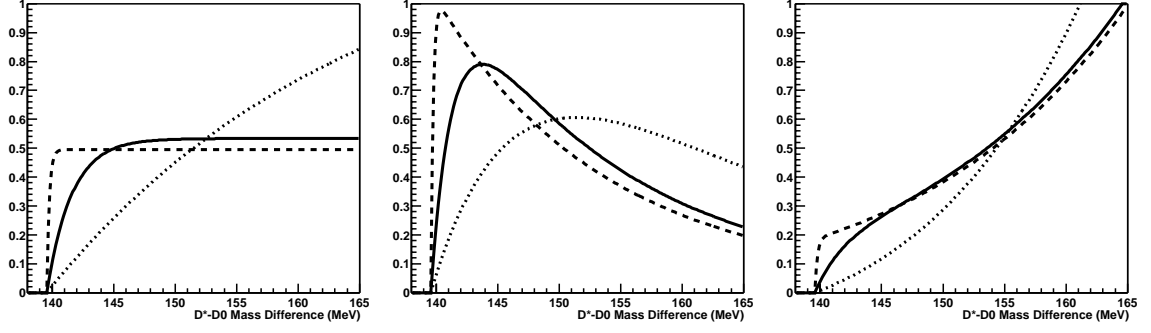
Figure 7.1: Effects of the combinatoric model shape parameters c_1 and c_2. Each plot is calculated using a different value of c_2: 0 (left), -10 (middle), +10 (right). The curves within each plot represent different values of c_1: 2 MeV (solid), 0.2 MeV (dashed), or 25 MeV (dotted). [25]

In order to fit for the signal and background shape parameters, the δm distribution is fit with the following PDF:

$$F(\delta m) = f_{peak} \cdot F_{peak}(\delta m) + (1 - f_{peak}) \cdot F_{comb}(\delta m) \tag{7.5}$$

The value of f_{peak} corresponds to the fraction of peak events in the whole sample.

Through extensive examination, it has been determined that the peak shape varies significantly inside two samples: SVT-DCH and D^0mode. The peak shape group(6) = SVTDCH(2)$\times D^0$mode(3) covers all possible peak shapes. Each peak group is broken down further into 12 subgroups which share a common set of 2 background shape parameters. The 12 groups are defined in table 7.3.

The 60 possible groups (TagCat(5)$\times$AngCut(2)$\times$LeptonID(3)$\times$OnOffRes(2)) are consolidated into 12 by making all Off-resonance events into one category, all Same-Side events into one category and both electrons and muons are lumped together so the 10 remaining categories are TagCat(5)$\times$(good/fake lepton).

For each peak group, there a 5 signal parameters + 12$\times$2 background parameters to find. A simultaneous fit is performed for each subgroup of the peak category by using

the likelihood function:

$$\mathcal{L} = \prod_{k=1}^{12} \prod_{events j} \left[f_{peak}^k \cdot F_{peak}(\delta m_j; m_1, \sigma_1, m_2, \sigma_2, f_1) + (1 - f_{peak}^k) \cdot F_{comb}(\delta m_j; c_1^k, c_2^k) \right]$$

(7.6)

where k labels the background sub-category. This fit introduces 12 fractions, f_{peak}^k, so the total number of floating parameters in the fit is $5 + 24 + 12 = 41$ [25].

BG Group	OnOffRes	AngCut	LeptonID	TagCat
Lept,LTag	On	OS	e,μ	Lepton Tag
Lept,KTag	On	OS	e,μ	Kaon Tag
Lept,NT1	On	OS	e,μ	NT1 Tag
Lept,NT2	On	OS	e,μ	NT2 Tag
Lept,NT3	On	OS	e,μ	NT3 Tag
Fake,LTag	On	OS	fake	Lepton Tag
Fake,KTag	On	OS	fake	Kaon Tag
Fake,NT1	On	OS	fake	NT1 Tag
Fake,NT2	On	OS	fake	NT2 Tag
Fake,NT3	On	OS	fake	NT3 Tag
SS	On	SS	-	-
OffRes	Off	-	-	-

Table 7.3: Breakdown of each peak group (which shares a common set of the 5 peak shape parameters) into separate background groups that each share a common set of the combinatoric background shape parameters c_1 and c_2 [25].

7.3 Peaking Backgrounds

Backgrounds other than combinatoric D^*'s are created with a good D^*, thus these backgrounds (peaking backgrounds) have the same δm shape as signal events. In order to determine the contributions of peaking backgrounds, the background fraction can be estimated using various control samples. The peaking background shapes are shown in figure 7.2.
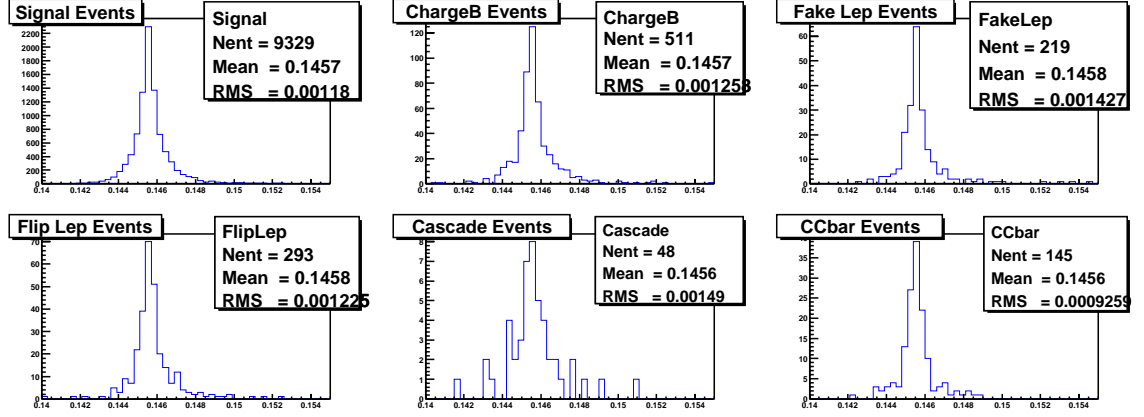
Figure 7.2: Signal δm shape compared to peaking backgrounds. The close similarity in mean and width of these distributions confirms the assumption that peaking background have the same δm shape as signal events. These plots were made with $\approx 10\,fb^{-1}$ of generic MC data and use the MC truth associator to pick out each particular background.

7.3.1 Fake Lepton

The number of observed events in the peak region of the electron, muon and fake lepton control samples (N_e, N_μ, N_f) can be related to the number of $D^*l\nu$ candidates with a real lepton or hadron through the following equations:

$$
\begin{pmatrix} N_e \\ N_\mu \\ N_f \end{pmatrix} = \begin{pmatrix} \epsilon_e(e\overline{\mu}) & \epsilon_\mu(e\overline{\mu}) & \epsilon_h(e\overline{\mu}) \\ \epsilon_e(\overline{e}\mu) & \epsilon_\mu(\overline{e}\mu) & \epsilon_h(\overline{e}\mu) \\ \epsilon_e(\overline{e\mu}) & \epsilon_\mu(\overline{e\mu}) & \epsilon_h(\overline{e\mu}) \end{pmatrix} \begin{pmatrix} \eta_e \\ \eta_\mu \\ \eta_h \end{pmatrix}
\tag{7.7}
$$

where $\epsilon_e(e\overline{\mu})$ is the efficiency for finding an electron given that the particle is a real electron, $\epsilon_h(\overline{e\mu})$ is the efficiency for finding a hadron given that the particle isn't an electron or muon. The diagonal terms should be close to one while the off diagonal terms correspond to various fake rates. The information used for the table is gathered from the Electron and Muon PID groups. The values are weighted averages based on the expected momentum distribution of events in the $D^*l\nu$ sample [25].

The true fractions of good electron, good muon and fake lepton events can be

determined by inverting the matrix and solving for each fraction.

7.3.2 Continuum Background

The continuum contribution can be determined by using off-peak data. Once the combinatoric and fake lepton contributions have been removed from the off-peak sample, the remaining events have a good D^* and good lepton. The number of continuum events is then scaled by the ratio of the on-peak/off-peak luminosities to determine the amount of continuum events in the signal data sample.

7.3.3 Uncorrelated Lepton

Once all the combinatoric, fake lepton and continuum events have been accounted for in the signal sample, the only events remaining are $B^0\overline{B}^0$ and B^+B^- events which produced a good D^* and a good lepton (not necessarily from the same B^0). A procedure very similar to the fake lepton case is used to determine the uncorrelated fraction. The number of peak events passing SS and OS cuts in the signal and uncorrelated samples can be related to each other with [25]:

$$\begin{pmatrix} N_{OS} \\ N_{SS} \end{pmatrix} = \begin{pmatrix} \epsilon_{OS}^{sig} & \epsilon_{OS}^{uncor} \\ \epsilon_{SS}^{sig} & \epsilon_{SS}^{uncor} \end{pmatrix} \begin{pmatrix} \eta_{sig} \\ \eta_{uncor} \end{pmatrix} \tag{7.8}$$

where the ϵ's are determined from unfiltered generic Monte Carlo.

The efficiencies depend on the Monte Carlo getting the relative efficiencies between these two samples correct, not absolute efficiencies. This is the only place in the entire measurement where Monte Carlo data is used to determine any parameter or value used to the actual data fit. The Monte Carlo is considered quite reliable for this task, and given the very small fraction of events in the flip lepton sample, any systematic error

introduced will be small compared to other expected systematic errors.

Once again the fractions can be determined by inverting the matrix. Table 7.4 shows the background fractions determined for the current dataset. Appendix B contains all relevant reference plots related to the determination of δm parameters and yields.

Category			Peak Yield	$f_{\text{cont.}}(\%)$	$f_{\text{fake}}(\%)$	$f_{\text{uncorr.}}(\%)$	$f_{\text{comb.}}(\%)$
On	OS	e	7008 ± 91	1.53 ± 0.42	0.168 ± 0.004	3.142 ± 0.011	17.89 ± 0.24
		μ	6569 ± 88	2.27 ± 0.57	2.669 ± 0.067	2.85 ± 0.25	18.36 ± 0.25
		f	8770 ± 108	12.8 ± 1.3	72.4 ± 1.8	0.7 ± 1.4	31.40 ± 0.24
	SS	e	306 ± 21	0.0 ± 0.006	0.533 ± 0.039	56.92 ± 0.20	34.0 ± 1.3
		μ	299 ± 20	5.1 ± 3.6	8.89 ± 0.64	48.9 ± 3.7	34.4 ± 1.3
		f	1350 ± 45	20.4 ± 4.1	74.4 ± 5.4	3.6 ± 7.0	42.59 ± 0.61
Off	OS	e	13.6 ± 3.8	100	0	0	21.0 ± 4.5
		μ	18.9 ± 4.7	100	0	0	25.1 ± 4.3
		f	142 ± 14	100	0	0	44.2 ± 1.8
	SS	e	0.0 ± 0.002	100	0	0	100 ± 100
		μ	1.9 ± 1.4	100	0	0	38 ± 14
		f	34.9 ± 7.1	100	0	0	52.5 ± 3.6

Table 7.4: Peak yields and fake, continuum and uncorrelated fractions of the peak yield, and the combinatoric fraction of total events in δm window for the various subsamples shown in Figs. B.8 and B.9. Peak yields include the peaking backgrounds [25]. In the table above, the sum of all peaking background plus signal fraction will equal one ($f_{\text{fake}} + f_{\text{continuum}} + f_{\text{uncorr.}} + f_{\text{signal}} = 1$). The last column describes the fraction of combinatoric background events, the fraction of peaking events is simply $1 - f_{\text{comb.}}$.

7.4 Time distribution PDF, $G(\Delta t, \sigma_{\Delta t}, g; \vec{q})$

7.4.1 Signal Δt PDF

The Δt PDF for correctly reconstructed events was introduced in the tagging chapter, equation 6.2. Section 6.6.1 describes a possible difference which should be observed between events which are tagged with a B^0 decay or a $\overline{B}^0$. This difference is accounted in our model by measuring the average mistag rate per tagging category and a Δmistag rate which describes the difference due to tagging flavor on the mistag rate.

The PDF is rewritten here in terms of ω (mistag rate) instead of D (dilution) and the Δt PDF is convoluted with a resolution function described in section 5.6 to account for detector resolution effects:

$$G(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,sig}) = \frac{1}{4\tau}e^{-|\Delta t|/\tau}(1+g(1-2(\omega\pm\Delta\omega/2))\cos(\Delta m\Delta t))\otimes R_{GExp+G}(\Delta t, \sigma_{\Delta t}; \vec{r}^{\,sig})$$

$$(7.9)$$

where $\vec{r}^{\,sig} = (\tau_{eff}, s, f, b_{out}, s_{out}, f_{out})$, $\vec{q}^{\,sig} = (\Delta m, \tau, \omega, \Delta\omega, \vec{r}^{\,sig})$. Also, $\omega = (\omega_{B^0}+\omega_{\overline{B}^0})/2$ and $\Delta\omega = \omega_{B^0} - \omega_{\overline{B}^0}$.

7.4.2 Background Δt PDF

The Δt PDFs for backgrounds can be composed of mixing, lifetime and prompt (zero lifetime) terms convoluted with a resolution function.

$$G_{mixing}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,mixing}) = \frac{1}{4\tau}e^{-|\Delta t|/\tau_{eff}}(1 + g(1 - 2\omega_{eff})\cos(\Delta m_{eff}\Delta t)) \qquad (7.10)$$

$$G_{lifetime}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,life}) = \frac{1}{4\tau}e^{-|\Delta t|/\tau_{life}}(1 + g(1 - 2\omega_{life})) \qquad (7.11)$$

$$G_{prompt}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,prompt}) = \frac{1}{2}(1 + g(1 - 2\omega_{prompt}))\delta(\Delta t) \qquad (7.12)$$

These PDF's are normalized correctly when both mixed and unmixed events are used. The mistag terms will cancel out leaving just the exponent term (equation 7.10, 7.11) or the $\delta(\Delta t)$ term (equation 7.12). The combinatoric background PDF, for example, is given by:

$$G_{background}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,back}) = (f \cdot G_{mixing} + (1 - f) \cdot G_{prompt}) \otimes R_{G+G}(\Delta t, \sigma_{\Delta t}; \vec{r}^{\,back})$$

$$(7.13)$$

where $\vec{r}^{\,back} = (b_1, s_1, b_{out}, s_{out}, f_{out})$, $\vec{q}^{\,back} = (\Delta m_{eff}, \tau_{eff}, \omega_{eff}, \omega_{prompt}, \vec{r}^{\,mixing})$.

The prompt term is intended to model the contribution of $c\bar{c}$ events, misreconstructed and flip lepton events where the reconstructed vertices are generally quite close

due to the lepton candidate being required to vertex with the D^*. The mixing term captures events that were almost correctly reconstructed. Many of the combinatoric D^*'s have the correct soft pion and kaon from the D^0, but one of the D^0 pion daughters is incorrect. These events still carry some information about B^0 decays and show mixing (but with some effective lifetime and mixing frequency which is not the same as signal events).

7.5 Asymmetry Plot

An asymmetry variable can be defined which very clearly shows the oscillating nature of $B^0\overline{B}^0$ events. The PDFs for unmixed/mixed events without Δt resolution ($f_\pm$) is given by equation 6.2: $f_\pm \sim e^{-|\Delta t|/\tau}(1 \pm D\mathbf{cos}(\Delta m\Delta t))$. The asymmetry variable is defined as:

$$A_{asym} = \frac{f_+ - f_-}{f_+ + f_-} = D\mathbf{cos}(\Delta m\Delta t) \tag{7.14}$$

A plot of A_{asym} shows the Dilution as the amplitude of the cos function and Δm_d as the frequency. This interpretation must be modified once detector resolution effects are added. The resolution function accounts most of the detector resolution effects, but now the amplitude of A_{asym} is no longer purely due to Dilution.

Figure 7.3 shows Δt distributions for unmixed events, mixed events and A_{asym}. The first plot shows distributions with perfect tagging (D=1) and perfect resolution. The remaining plots show the effects of imperfect tagging and resolution separately and combined. The "Perfect Dt" plot reveals the dramatic effect mistags have on the mixed sample. The mistag rate in this sample is $\approx 10\%$. The unmixed sample is much larger than the mixed sample, so 10% of unmixed events migrating to the mixed sample has a very

large effect. The "Perfect Tag" plot shows how resolution effects can alter the magnitude

of A_{asym}. Even with perfect tagging, the magnitude of A_{asym} will be less than 1 because

by including resolution effects, the $\Delta t = 0$ bin now contains events. The final plot shows

the combined effects of imperfect tagging and resolution.
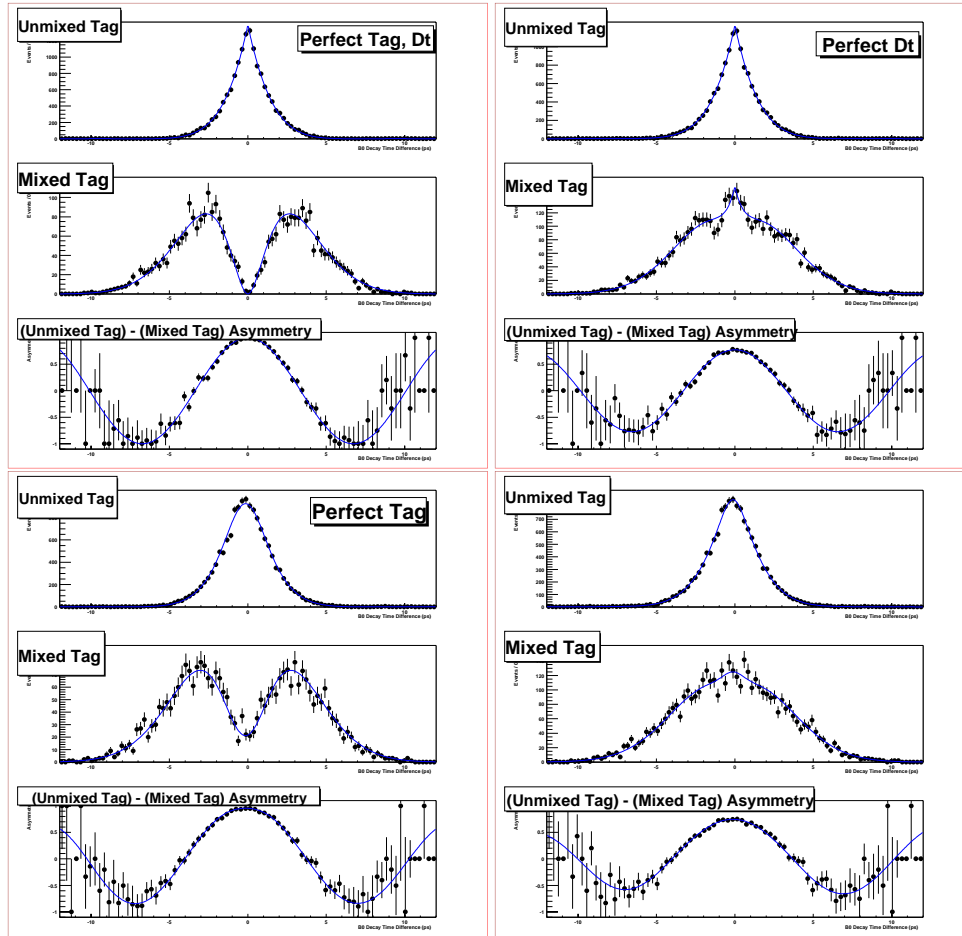
Figure 7.3: Effects of imperfect tagging and Δt on the mixed, unmixed and A_{asym} shapes.

7.6 Fitting Strategy

The final fit is performed as a simultaneous fit over all 12 signal and control samples (AngCut(2)×OnOffRes(2)×LeptonID(3) - indexed by s) which are subdivided into 30 subsamples (TagCat(5)×D0mode(3)×SVTDCH(2) - indexed by c) using the likelihood [23]

$$\mathcal{L} = \prod_{s=1}^{12} \prod_{c=1}^{30} \prod_{k=1}^{N(s,c)} P_{s,c}(\vec{x}_k; \vec{a}) \tag{7.15}$$

where k indexes the N(s,c) events in each of the 360 subsamples. The PDF used in the likelihood function is given by [23]:

$$P_{s,c}(\delta m, \Delta t, \sigma_{\Delta t}, g; \vec{a}) =$$

$$f_{s,c}^{comb} \cdot F^{comb}(\delta m; \vec{p}_{s,c}^{\,comb}) \cdot G^{comb}(\Delta t, \sigma_{\Delta t}, g; \vec{q}_{s,c}^{\,comb}) + (1 - f^{comb}) \cdot F^{peak}(\delta m; \vec{p}_c^{\,peak}) \cdot$$

$$\left[\sum_{j=1}^{3} f_{s,c,j}^{peak} \cdot G_j^{peak}(\Delta t, \sigma_{\Delta t}, g; \vec{q}_{s,c,j}^{\,peak}) + \left(1 - \sum_{j=1}^{3} f_{s,c,j}^{peak} \right) \cdot G^{sig}(\Delta t, \sigma_{\Delta t}, g; \vec{q}_c^{\,sig}) \right] \tag{7.16}$$

See table 7.5 for a short description of each term in $P_{s,c}$. The equation for $P_{s,c}$ has two explicit assumptions: (1) the peaking backgrounds have the same δm shape as signal events and (2) the Δt signal parameters (q_c) are indexed only by c, which means that the $\Delta t, \sigma_{\Delta t}$ shapes of signal events are the same whether the event is in the signal sample or a control sample (i.e. a real signal event in the fake or flip lepton samples will have the same Δm_d and lifetime as events in the signal sample).

The parameters in the full PDF given by equation 7.16 are not broken down by all 360 possible different sub-categories. A systematic Monte Carlo study is completed for each parameter to determine what categories the parameter actually depends on. The categories are completely flexible and can be tailored to any specification. The next chapter on Monte Carlo validation details this process.

Ideally, the fit would be performed by allowing all parameters to float. Floating

all parameters is the most reliable way to determine correlations between parameters and the systematic errors in the fit. This is not practical for two reasons, (1) there are too many parameters and the fit would take an extremely long time, (2) some parameters can't be fitted, they are determined from external inputs like lepton mis-id rate. The approached used for this analysis is a sequence of fits which pin down different sets of parameters. This will reduce the total number parameters floating in the final fit.

The parameters used for the signal probability function, $F(\delta m; \vec{p})$, and the combinatoric portion of the Δt distribution function, $G^{comb}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,comb})$, can be determined without approximations or simplifications. The Δt peaking background parameters, however, can't be independently fit because of signal, combinatoric and other type of events contaminating the control samples. In order to overcome these issues, the following series of fits and calculations are performed.

1. δm **fit**: Simultaneous fit for peaking and background shapes using equation 7.6. Determines $\vec{p}^{\,comb}_{s,c}$ and $\vec{p}^{\,peak}_{c}$

2. δm **yield**: Calculate yield fits for each subsample (s,c). Determines $f^{\,comb}_{s,c}$.

3. **Background fractions**: Calculate the peaking background fractions using externally derived quantities, sections 7.3.1, 7.3.2 and 7.3.3. Determines $f^{\,peak}_{s,c,p}$.

4. **Combinatoric Δt fit**: Fit for combinatoric Δt parameters using the combinatoric sideband (0.155< δm <0.165 GeV). The sideband is essentially 100% combinatoric events so this fit requires no information about the signal Δt shape. Once this step is done, the values of $F(\delta m; \vec{p})$ and $G^{comb}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,comb})$ have been completely determined. These functions are precalculated for later fits.

5. **Bootstrap signal fit**: Each of the peaking background fits require some knowledge

of the signal Δt shape. The strategy used here is to fit for peaking Δt shape parameters and uses these as an approximation to the signal Δt shape parameters.

Signal events dominate the signal region, so deviations between the peaking Δt shape and signal Δt shape should be small. Signal events are generally a small contamination in the background control samples so these deviations will be a second order effect in the control sample fits.

6. **Continuum fit**: The ccbar contribution is determined by fitting the off-resonance data and assuming the only contributions are combinatoric events and ccbar events. Figure B.9 shows that this is a very good assumption. This fit determines $\vec{q}_{s,c,1}^{\,peak}$.

7. **Fake lepton fit**: The fake lepton sample is fairly large. This sample is fit assuming that there is only combinatoric, fake lepton, ccbar and signal events. Figure B.8 shows how small the other backgrounds are. This fit determines $\vec{q}_{s,c,2}^{\,peak}$.

8. **Uncorrelated fit**: The uncorrelated lepton fit is the final peaking background fit. This fit determines $\vec{q}_{s,c,3}^{\,peak}$.

9. **Final fit**: Now that all the peaking background Δt parameters are determined, the final fit is performed with the signal Δt parameters, $\vec{q}_c^{\,sig}$, floating. This final fit yields the B^0 lifetime, Δm_d and mistag rates plus resolution parameters for the signal events.

Function	Description
P	Overall probability to observe an event
F^{comb}	Probability of observing δm for a combinatoric BG event
F^{peak}	Probability of observing δm for an event with a real D^*
G^{comb}	Probability of observing Δt, $\sigma_{\Delta t}$ for a combinatoric BG event
G_1^{peak}	Probability of observing Δt, $\sigma_{\Delta t}$ for a continuum BG event
G_2^{peak}	Probability of observing Δt, $\sigma_{\Delta t}$ for a fake-lepton BG event
G_3^{peak}	Probability of observing Δt, $\sigma_{\Delta t}$ for an uncorrelated $D^* l\nu$ BG event
G^{sig}	Probability of observing Δt, $\sigma_{\Delta t}$ for a signal event
$f_{s,c}^{comb}$	Fraction of combinatoric events in subsample (s,c) (normalized to all subsample events in the δm window)
$f_{s,c,j}^{peak}$	Fraction of peaking BG events of type j in subsample (s,c) (normalized to subsample events in the δm peak)
$\vec{p}_{s,c}^{comb}$	Combinatoric BG δm shape parameters for subsample (s,c)
$\vec{p}_c^{peak}$	δm Peak shape parameters for subsample (c)
$\vec{q}_{s,c}^{comb}$	Combinatoric BG $\Delta t, \sigma_{\Delta t}$ shape parameters for subsample (s, c)
$\vec{q}_{s,c,j}^{peak}$	Peaking BG type-p $\Delta t, \sigma_{\Delta t}$ shape parameters for subsample (s,c)
$\vec{q}_c^{sig}$	Signal $\Delta t, \sigma_{\Delta t}$ shape parameters for subsample (c)

Table 7.5: Description of the functions appearing in the likelihood equation 7.15. Functions labeled with F calculate a probability of observing a particular value of δm while functions labeled with G give probabilities for values of Δt and $\sigma_{\Delta t}$. Parameters labeled with f describe the relative contributions of different types of events. Parameters labeled with $\vec{p}$ describe the shape of a δm distribution, and those labeled with $\vec{q}$ describe a Δt, $\sigma_{\Delta t}$ shape.

Chapter 8

Monte Carlo Validation

This chapter describes many of the studies and fits used to validate the fitting procedure and study systematics biases and errors. The fits for signal and backgrounds are done separately with the help of Monte Carlo truth information to see if the individual fits make sense. Different backgrounds are added to the signal one at a time to build up to the full fitting sequence.

8.1   Signal Δt fit

This section looks at the Δt fit for correctly reconstructed events. This is accomplished by applying all standard $D^* l \nu$ cuts to the signal Monte Carlo sample plus a cut requiring the events to be MC matched signal events. The PDF used for the fit is the full PDF, equation 7.15, with f^{comb}=0, f^{peak}=0 and F^{peak}=1 or simply $P_{s,c}(\delta m, \Delta t, \sigma_{\Delta t}, g; \vec{a}) = G^{sig}(\Delta t, \sigma_{\Delta t}, g; \vec{q}_c^{\,sig})$.

Table 8.1 shows the values for Δm_d and lifetime using Monte Carlo truth vertexing and tagging information. The mistag rates are calculated with a simple counting of events correctly tagged in each category. Table 8.2 shows the fits without using Monte

Carlo truth information.

τ_{B^0} (ps)	Δm_d (ps^{-1})	ω_L	ω_K	ω_{NT1}	ω_{NT2}	ω_{NT3}
1.5446 ± 0.0051	0.4708 ± 0.0013	6.3	15.7	19.2	33.8	46.5

Table 8.1: τ_{B^0} and Δm_d are from fit to the true Δt and true tagging information. Mistag fractions are from counting wrong tagging events according to the MC truth. The mistag rates are quoted in percent.

The first important question to address is what parameters will be broken down into what categories. For any Δt fit, there is no reason to believe that the decay mode or SVT/SVT+DCH categories affect the Δt structure. Signal events are assumed to have the same Δt structure in all the control samples as the signal sample, thus there is no need to consider the three categories referring to control samples (LeptonID, AngCut, OnOffRes). This leaves TagCat as the only category relevant for signal events. The value of Δm_d and τ will not be broken down by TagCat, we are looking for the global value of these parameters. The mistag rates and Δmistag rates clearly need to be broken down by TagCat. The only remaining question is what to do with the resolution function.

The outlier terms will be fit as global quantities, these terms are very insensitive to the different tagging categories. Figure 5.15 shows that the per-event-error removes the need to break the core scale factor down by TagCat, thus only the core bias, core fraction (3G resolution model)/effective lifetime,core fraction (GExp resolution model) will be broken down by TagCat.

The signal fit will use the fact that Kaon mistag rate scales with $\sigma_{\Delta t}$ as shown in section 6.7. The mistag rate for all categories will be defined as an offset term plus a slope times the $\sigma_{\Delta t}$: $\omega = \omega_{offset} + \omega_{slope} \cdot \sigma_{\Delta t}$. The slope for all categories except the Kaon is fixed to zero in all fits. Figure 8.1 shows the asymmetry plots for each tagging category.

	GExp				
	Lept	Kaon	NT1	NT2	NT3
τ			1.553 ± 0.009		
Δm_d			0.473 ± 0.004		
$\Delta\omega$	0.004 ± 0.006	-0.014 ± 0.005	0.012 ± 0.010	-0.042 ± 0.008	-0.074 ± 0.006
ω_{off}	0.056 ± 0.003	0.068 ± 0.008	0.184 ± 0.006	0.339 ± 0.005	0.462 ± 0.004
ω_{slope}	0	0.127 ± 0.012	0	0	0
f_{core}	0.822 ± 0.027	0.645 ± 0.028	0.838 ± 0.034	0.711 ± 0.030	0.685 ± 0.025
τ_{eff}			1.014 ± 0.078		
s_{core}			0.999 ± 0.022		
f_{out}			0.012 ± 0.003		
μ_{out}			-1.66 ± 0.67		
s_{out}			5.36		
	G+G+G				
τ			1.551 ± 0.009		
Δm_d			0.475 ± 0.004		
$\Delta\omega$	0.004 ± 0.006	-0.014 ± 0.005	0.012 ± 0.010	-0.042 ± 0.008	-0.074 ± 0.006
ω_{off}	0.055 ± 0.003	0.068 ± 0.009	0.183 ± 0.006	0.339 ± 0.005	0.462 ± 0.004
ω_{slope}	0	0.126 ± 0.012	0	0	0
μ_{core}	-0.067 ± 0.034	-0.245 ± 0.025	-0.043 ± 0.04	-0.175 ± 0.032	-0.193 ± 0.026
s_{core}			1.051 ± 0.024		
f_{core}			0.906 ± 0.020		
μ_{wide}			-1.477 ± 0.298		
s_{wide}			2.397 ± 0.220		
f_{out}			0.004 ± 0.002		
μ_{out}			-2.73 ± 1.50		
s_{out}			8		

Table 8.2: Simultaneous fit for parameters using the GExp+G and G+G+G resolution model. The data set is all correctly reconstructed $D^*l\nu$ events from the signal Monte Carlo sample. The outlier scale factors are fixed to improve fit stability.
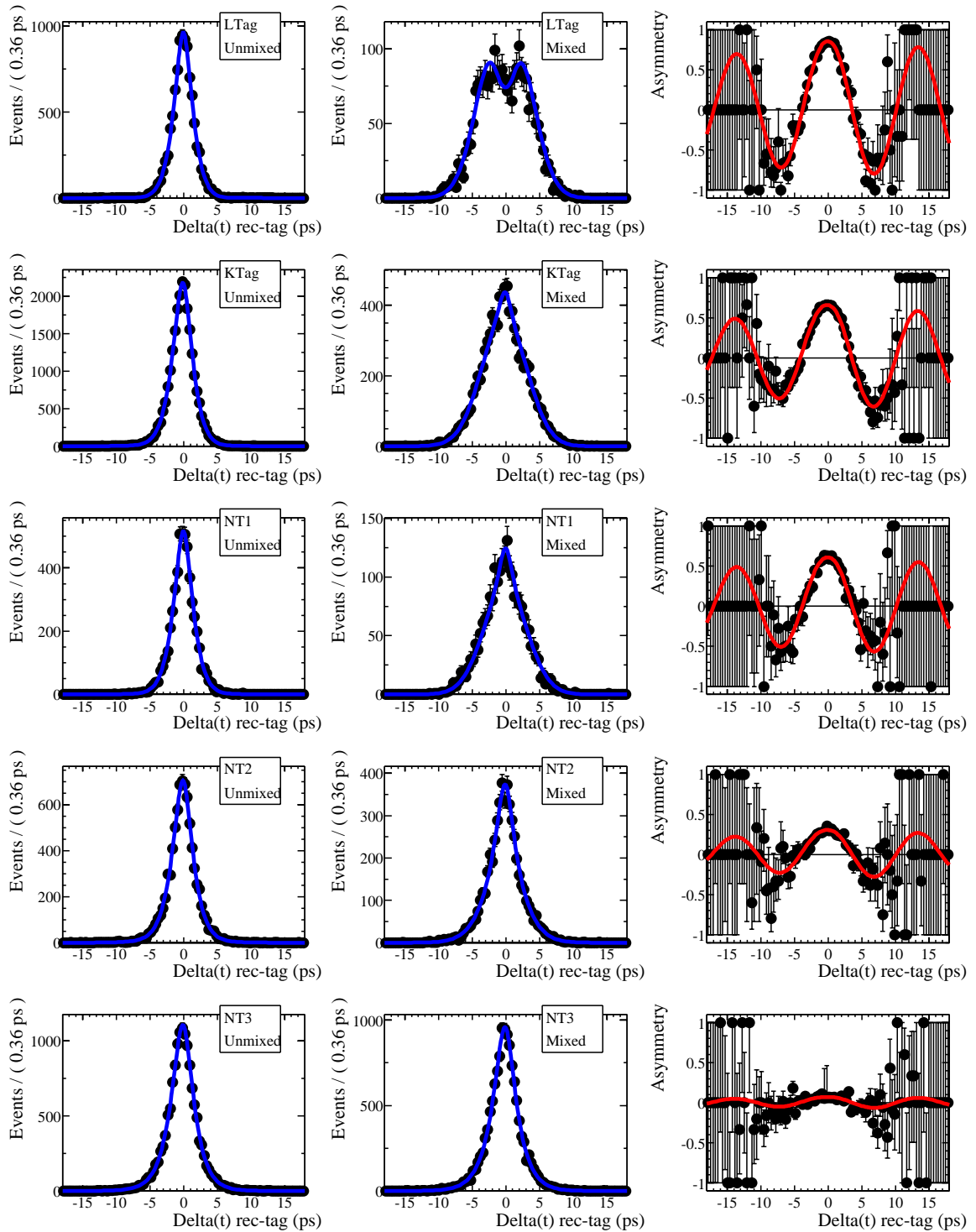
Figure 8.1: Δt distributions of unmixed, mixed and asymmetry plots for five tagging categories. The differing oscillation amplitudes show the large difference in tagging power of the various tagging categories.

8.2   Combinatoric Δt fit

The combinatoric background is the largest background in the $D^*l\nu$ sample. This type of background results from one or more of a D^*'s daughter's being incorrectly identified. The control sample for combinatorics is the D^*-D^0 sideband window, in the range 0.155-0.165 GeV. A modest fraction ($\approx$25%) of these events are $D^*l\nu$ events which were incorrectly reconstructed, the rest came from a decay which was not $D^*l\nu$.
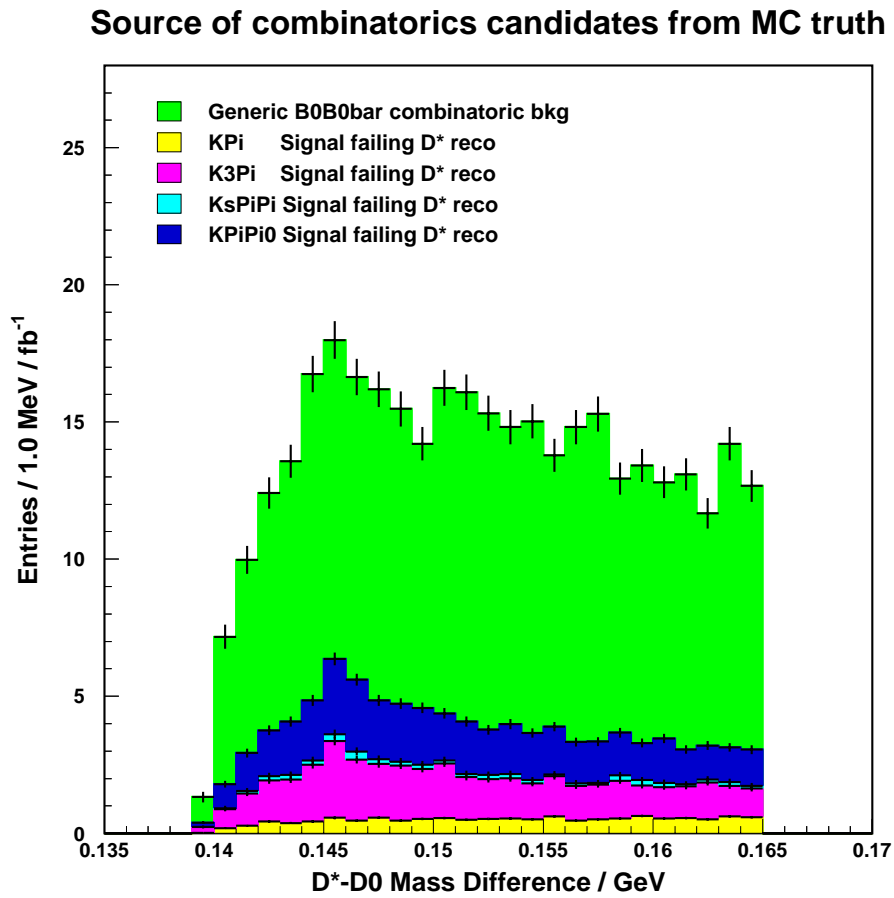
Figure 8.2: This figure compares the sources of combinatoric candidates using Monte Carlo truth association. The sum of the combinatoric candidates from the 4 signal modes represent only 25% of all those found in $B^0\overline{B}^0$ generic Monte Carlo. The signal and generic Monte Carlo are normalized so that they have the same total number of events (signal+background).

The PDF which describes the combinatoric background is given by equation 8.2. The categories used to break down the parameters in the PDF were discovered through an extensive iterative process of testing many different divisions and seeing how consistent the parameters are inside any given division. The final answer needs to balance maintaining a reasonable number of parameters vs finding all the possible category dependencies. The final categories used to break down the combinatoric background are given in table 8.3.

Parameter	Group	# pars
$bias_{core}$	$(L_{or}NT1_{or}NT2Tag) \oplus (K_{or}NT3)$	2
$scale_{core}$	$(OffRes_{or}SS) \oplus (OnRes_{and}OS)$	2
η_{osc}	$tagCatGrp \otimes (Lept \oplus Fake)$	10
f_{osc}	$OffRes \oplus SS \oplus TwoLept \oplus OneLept \oplus NoLept$	5
Δm_{BG}		1
τ_{BG}		1
η_{prompt}	common to all	1
$f_{outlier}$		1
$bias_{outlier}$		1
$scale_{outlier}$		1

Table 8.3: The final parameter breakdown by categories. Most breakdown categories have been previously defined. The TwoLept category is for events with a good lepton and a lepton tag, OneLept is an event with either a good lepton with no lepton tag or a fake lepton with a lepton tag, NoLept is an event in the fake sample with no lepton tag.

It is very important to know whether the events in the combinatoric sideband are representative of the events in the signal region. If the fit parameters are closely matched, then it is reasonable to measure the Δt structure of the sideband and use the measured parameters for characterizing the combinatorics under the signal peak.

Figure 8.3 in shows that the fraction of $B^0\overline{B}^0$ declines with an increasing value of δm. One could expect that the effective oscillation frequency could change as a function of δm. Since the fraction of charged B's is higher in the sideband, one would expect that the effective oscillation frequency is smaller in the sideband. A study was done

to determine how the combinatoric background parameters varied as a function of δm.

Figure 8.4 shows how the Δm_d parameters vary from the sideband into the signal region.

The best fit line has a small negative slope, but it is also consistent with a slope of zero.

Including a δm dependence for the combinatoric Δm_d parameter would be a significant complication, and since the effect is small - the effect will not be included and a systematic error will be estimated for this omission.

Figure 8.5 shows the total combinatoric fit and data projected onto 10 different plots (tagCatGrp⊗mixed/unmixed). The PDF is not explicitly broken down by this division, so the fact that the projections are all reasonable gives a good measure of how well the fit performs. Figure 8.6 shows the same for leptID⊗mixed/unmixed.

Source	Fraction under peak	Fraction in sideband
B^+B^-	0.216	0.376
$B^0\overline{B}^0$	0.760	0.598
$c\bar{c}$	0.024	0.026

Table 8.4: Relative composition of combinatoric backgrounds in the signal region ($143 < \delta m < 148$ MeV) and in the sideband region ($155 < \delta m < 165$ MeV).
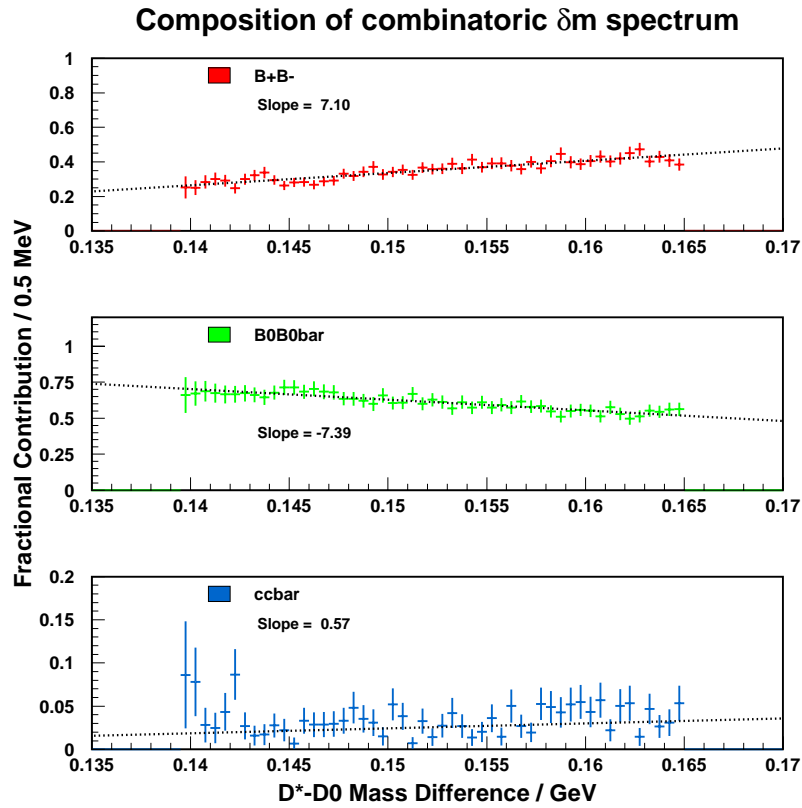
Figure 8.3: Relative composition of the combinatoric δm spectrum from Monte Carlo. Note that $B^0\overline{B}^0$ dominates everywhere, although the relative composition between $B^0\overline{B}^0$ and B^+B^- changes by 15% from the signal region to the far sideband.
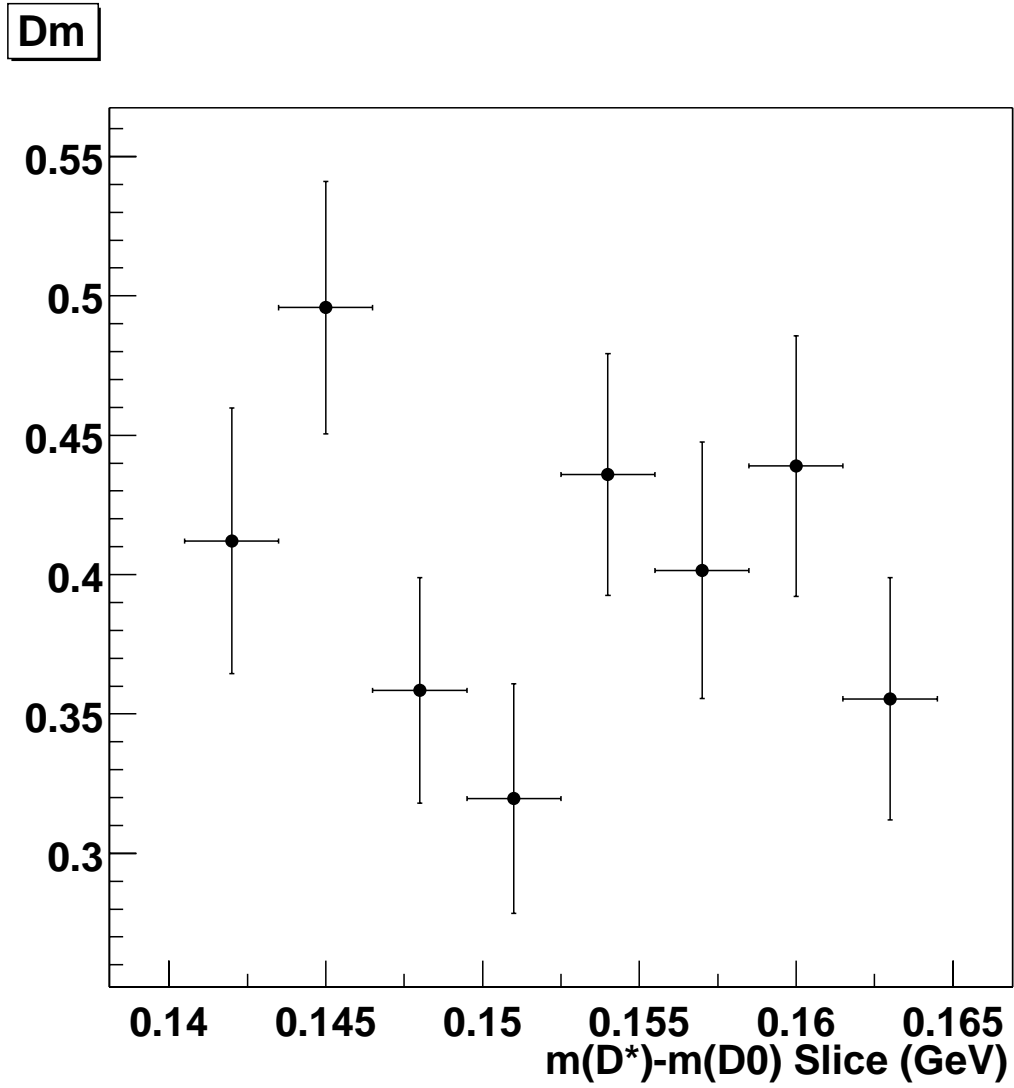
Figure 8.4: This plot shows the dependence of Δm_d on δm. The data points are consistent with a line with zero slope. This means that any δm dependence of Δm_d can be ignored at the cost of a small contribution to the total systematic error.

f_{TwoLept}	**1.000**
f_{OneLept}	0.887 ± 0.019
f_{NoLept}	0.700 ± 0.018
f_{OffRes}	0.159 ± 0.056
f_{SS}	0.485 ± 0.026
τ **(ps)**	1.257 ± 0.021
Δm **(ps^{-1})**	0.403 ± 0.021
μ_{LNT1NT2}	-0.064 ± 0.021
$\mu_{\text{K}-\text{NT3}}$	-0.154 ± 0.012
σ_{OnOS}	1.146 ± 0.024
σ_{OffSS}	1.247 ± 0.030
$\eta_{\text{o; LTag;Fake}}$	0.406 ± 0.023
$\eta_{\text{o; KTag;Fake}}$	0.192 ± 0.014
$\eta_{\text{o; NT1;Fake}}$	0.386 ± 0.027
$\eta_{\text{o; NT2;Fake}}$	0.455 ± 0.018
$\eta_{\text{o; NT3;Fake}}$	0.519 ± 0.014
$\eta_{\text{o; LTag;Lept}}$	0.243 ± 0.024
$\eta_{\text{o; KTag;Lept}}$	0.268 ± 0.010
$\eta_{\text{o; NT1;Lept}}$	0.276 ± 0.024
$\eta_{\text{o; NT2;Lept}}$	0.426 ± 0.016
$\eta_{\text{o; NT3;Lept}}$	0.488 ± 0.011
η_p	0.432 ± 0.012
f_{core}	0.981 ± 0.005
μ_{outlier}	-1.673 ± 0.508
σ_{outlier}	6.583 ± 0.693

Table 8.5: Parameters for a simultaneous fit to generic Monte Carlo data. The outlier scale and bias were fixed to globally determined quantities. The term f_{TwoLept} is extremely close to 1 in all fits and was fixed to 1 to avoid problems with fitted value being equal to the parameter limit.
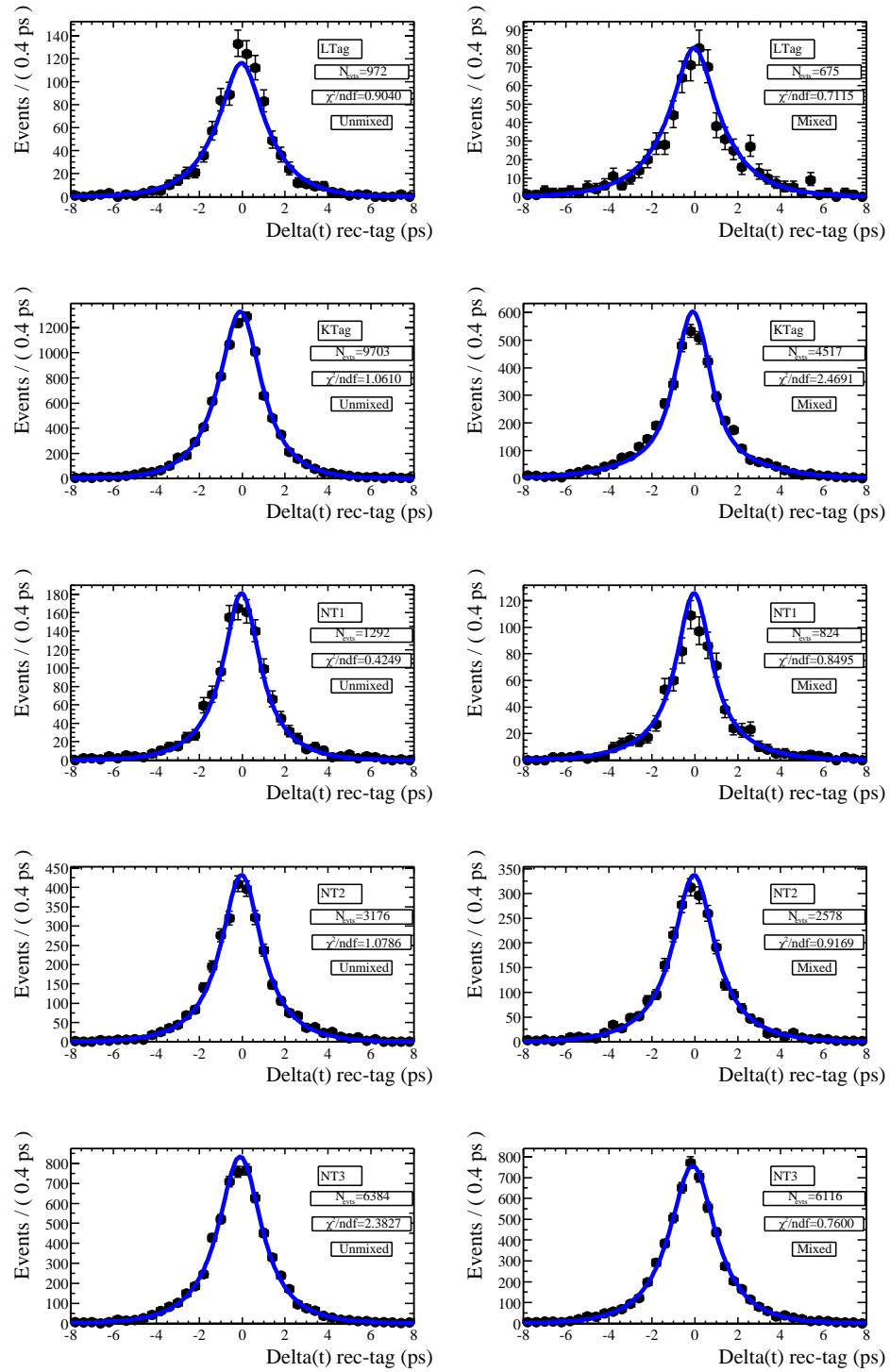
Figure 8.5: The combinatoric fit shown projected by mixing and tagCat status.
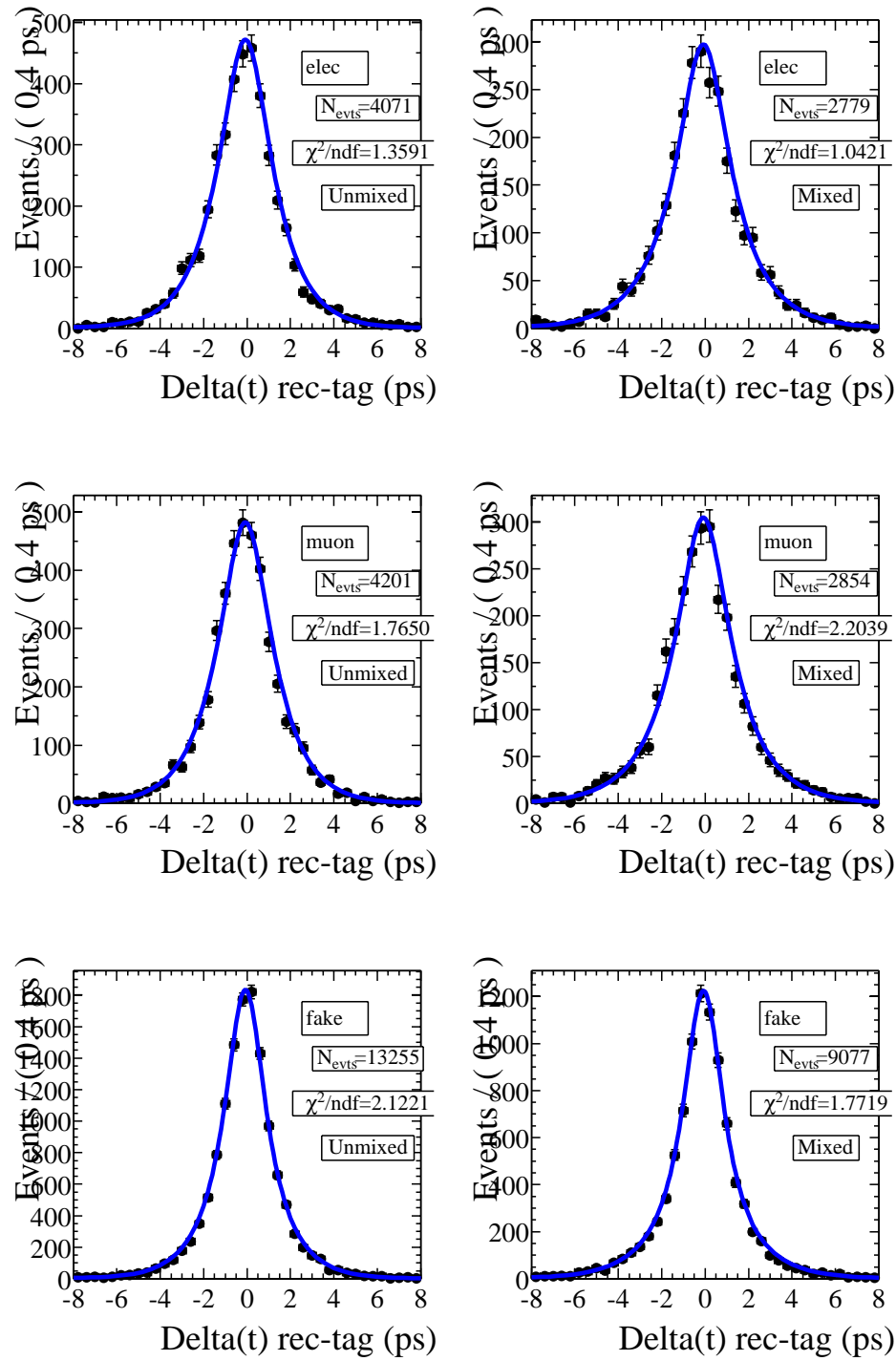
Figure 8.6: The combinatoric fit shown projected by mixing and leptID status.

8.3   Fake Lepton Δt fit

The fake lepton events are events which pass all signal region cuts except the lepton PID cut. These events are modeled with the same PDF as the combinatorics, a mixing plus prompt term convoluted with a double Gaussian resolution model. The fake lepton sample is fairly large so it is able to support a large number of parameters and parameter break down like the combinatoric sample. The parameter break down is shown in table 8.6.

Parameter	Group	# pars
$\mathrm{bias}_{\mathrm{core}}$	$(\mathrm{L_{or}NT1_{or}NT2Tag}) \oplus (\mathrm{Kaon_{or}NT3})$	2
η_{prompt}	$(\mathrm{L_{or}NT1_{or}NT2Tag}) \oplus (\mathrm{Kaon_{or}NT3})$	2
η_{osc}	tagCatGrp	5
Δm_{BG}		1
τ_{BG}		1
f_{prompt}		1
$\mathrm{scale}_{\mathrm{core}}$	common to all	1
f_{outlier}		1
$\mathrm{bias}_{\mathrm{outlier}}$		1
$\mathrm{scale}_{\mathrm{outlier}}$		1

Table 8.6: The final parameter breakdown by categories for the Fake Lepton sample.

The total number of parameters is 16, 6 of which are resolution parameters which will be combined with the other peaking backgrounds to form a common set of peaking resolution parameters.

One measure of the goodness of fit is shown in figure 8.7. The PDF fit and data are projected onto ten different plots (tagCatGrp$\otimes$mixed/unmixed). The PDF fits the data well in each projection, which shows good agreement since the PDF isn't intrinsically broken down by these projections but still fits them well.

Once again, since all Δt parameters will float in the final fit, an initial bootstrap

fit of the fake sample is only required to determine some reasonable starting points for the fake lepton parameter for the final fit.
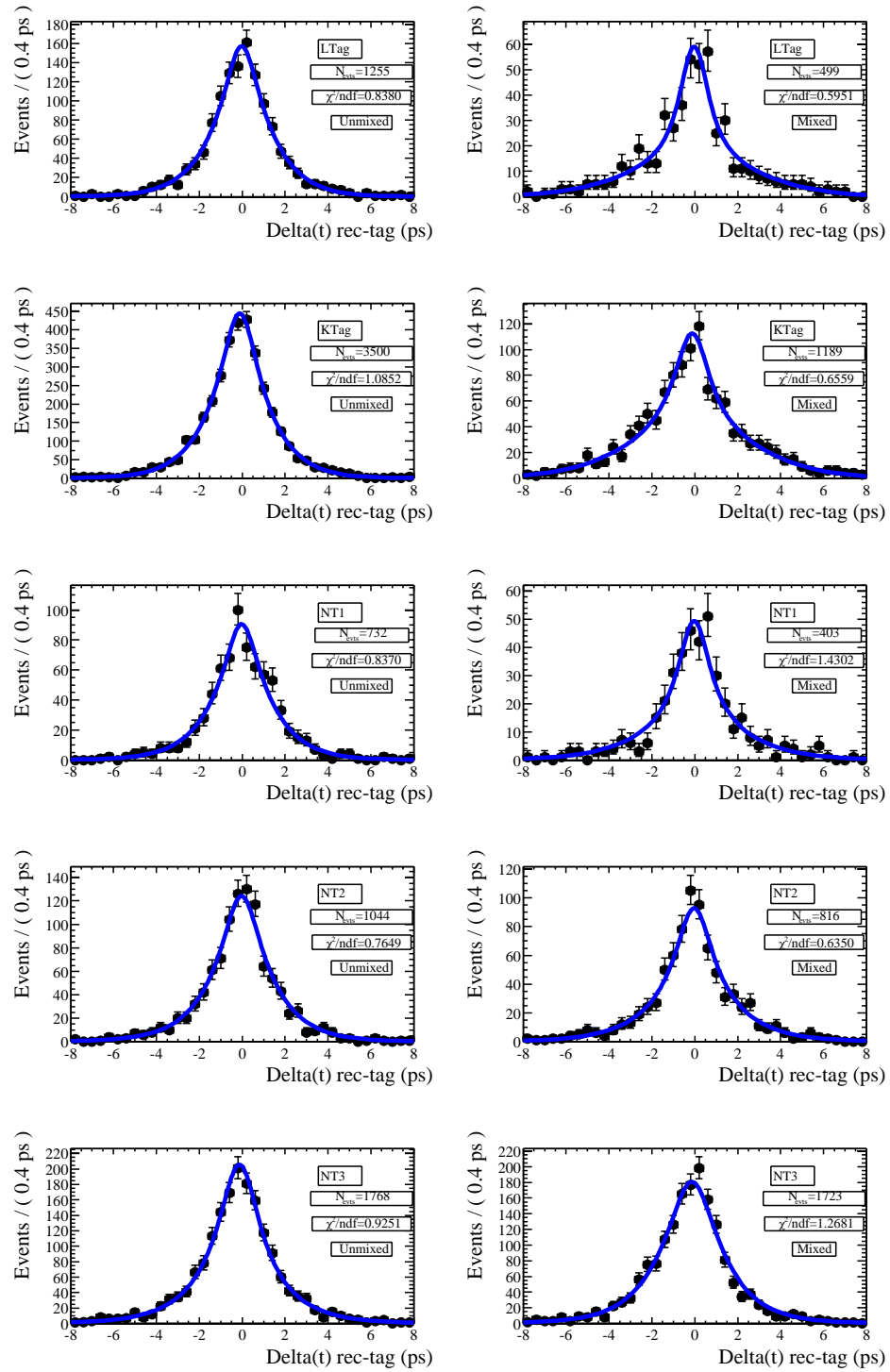
Figure 8.7: The fake lepton data and PDF projected onto (tagCatGrp⊗mixed/unmixed).
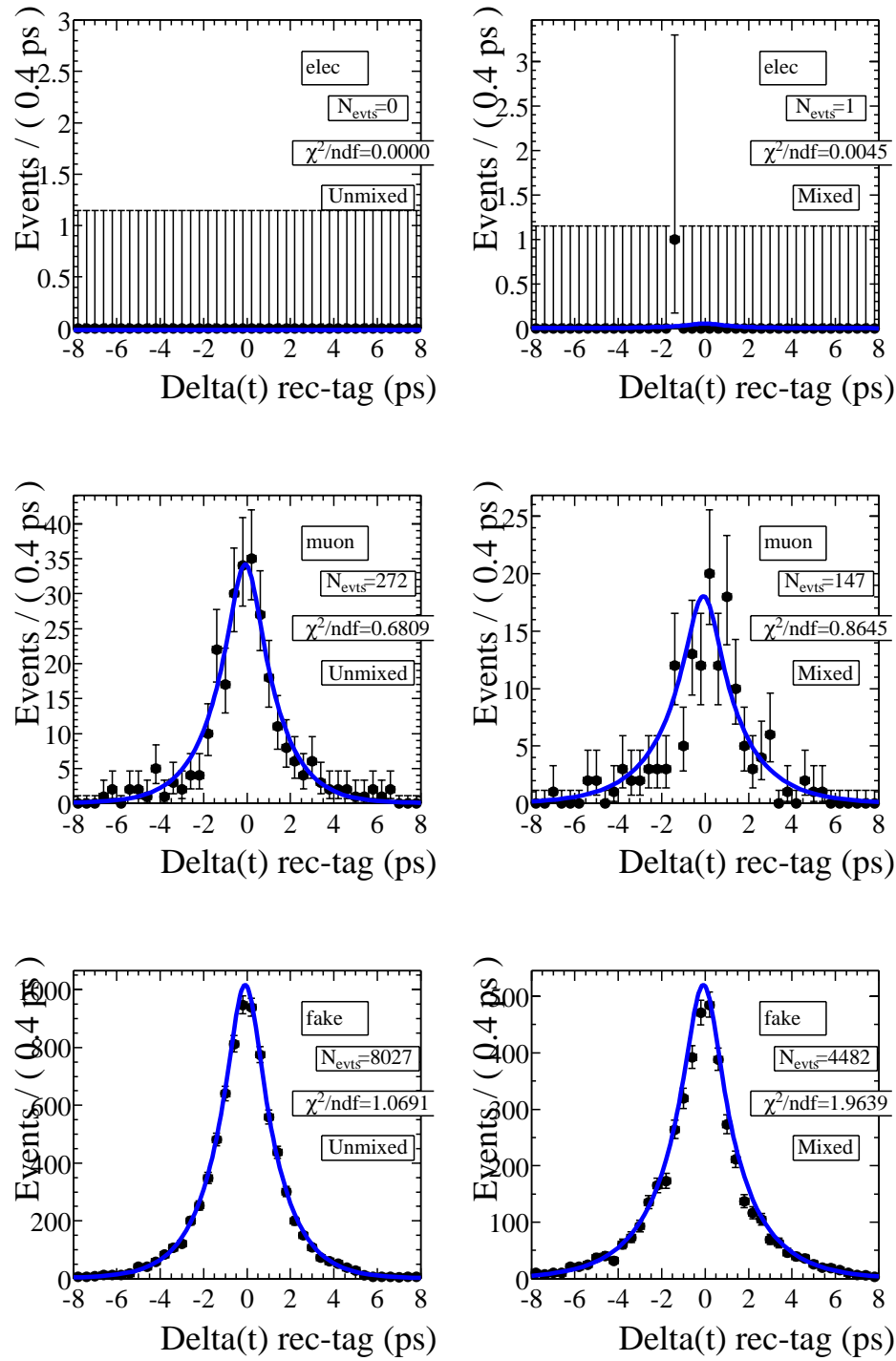
Figure 8.8: The fake lepton data and PDF projected onto (leptID⊗mixed/unmixed).

8.4 Uncorrelated Lepton Δt fit

The uncorrelated lepton sample contains flip lepton events, where the D^* comes from one B^0 while the real lepton comes from the other, and cascade events where the D^* and lepton ultimately come from the same B^0 parent but the lepton isn't the primary lepton. The flip leptons make up the vast majority of this background.

There are two physics considerations which help determine the form of the PDF used to model the uncorrelated leptons. First, events with a small value of true Δt have a better chance of being a flip lepton since the lepton and D^* must pass a vertex cut. Events with a large true Δt separation probably wouldn't pass such a cut. Also, the reconstructed Δt will probably be smaller than the true Δt since tracks from both B^0's are interchanged - drawing the reconstructed vertices closer. The end result is that the true Δt distribution of these events tends to be narrower than signal events and the reconstructed Δt are even more so. Thus a large prompt component for these types of events is expected.

The second consideration is that the true mixing status plays a role. Events which are mixed are more likely to produce flip lepton events due to the correct charge correlation of the lepton with the D^*. The loss of a high momentum lepton from the tag side and incorrect reconstruction hurts the tagging algorithm. Combine this with the fact that the uncorrelated lepton sample isn't very large and can't support too many physics parameters and the right answer is to combine the prompt component with a lifetime component convoluted with a single Gaussian resolution model.

$$G_{uncorr}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,uncorr}) = (f \cdot G_{life} + (1 - f) \cdot G_{prompt}) \otimes R_G(\Delta t, \sigma_{\Delta t}; \vec{r}^{\,uncorr}) \quad (8.1)$$

The performance of this model is tested by projecting it onto different subsam-
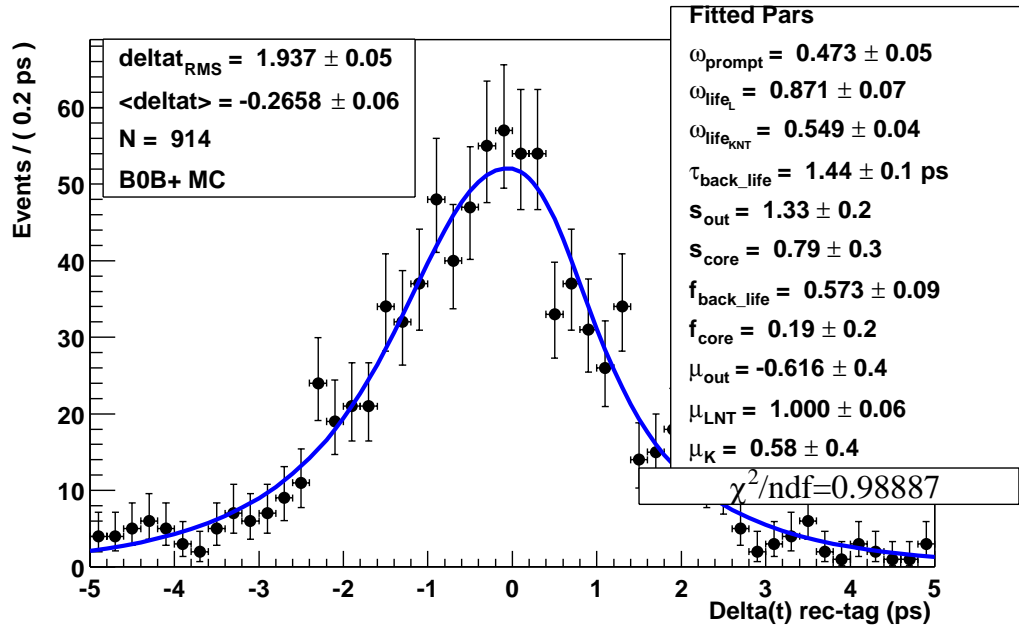
Figure 8.9: Δt projection of the simultaneous fit to the pure uncorrelated lepton sample drawn from generic Monte Carlo.

ples of the data. For instance, Fig. 8.10 shows that the model is generally consistent

subsamples for the data separated by lepton identification.

Parameter	Group	# pars
η_{life}	LTag$\otimes(Kaon_{or}NT1_{or}NT2_{or}NT3)$	2
τ_{BG}		1
η_{prompt}		1
f_{life}		1
$\text{bias}_{\text{core}}$	common to all	1
$\text{scale}_{\text{core}}$		1
f_{outlier}		1
$\text{bias}_{\text{outlier}}$		1
$\text{scale}_{\text{outlier}}$		1

Table 8.7: The final parameter breakdown for the Uncorrelated Lepton sample.

Parameter	Fitted Value
τ_{bkg} (ps)	1.22 ± 0.1
$\omega_{life_{KNT}}$	0.550 ± 0.04
ω_{life_L}	0.765 ± 0.07
ω_{prompt}	0.415 ± 0.08
f_{life}	0.55 ± 0.1
f_{core}	0.965 ± 0.03
μ	-0.330 ± 0.03
μ_{out}	-1.37 ± 2
σ	1.37 ± 0.1
σ	5.7 ± 2

Table 8.8: Parameter values for a model of prompt plus lifetime terms convolved with a single Gaussian resolution function fitted to the pure uncorrelated sample drawn from generic Monte Carlo.
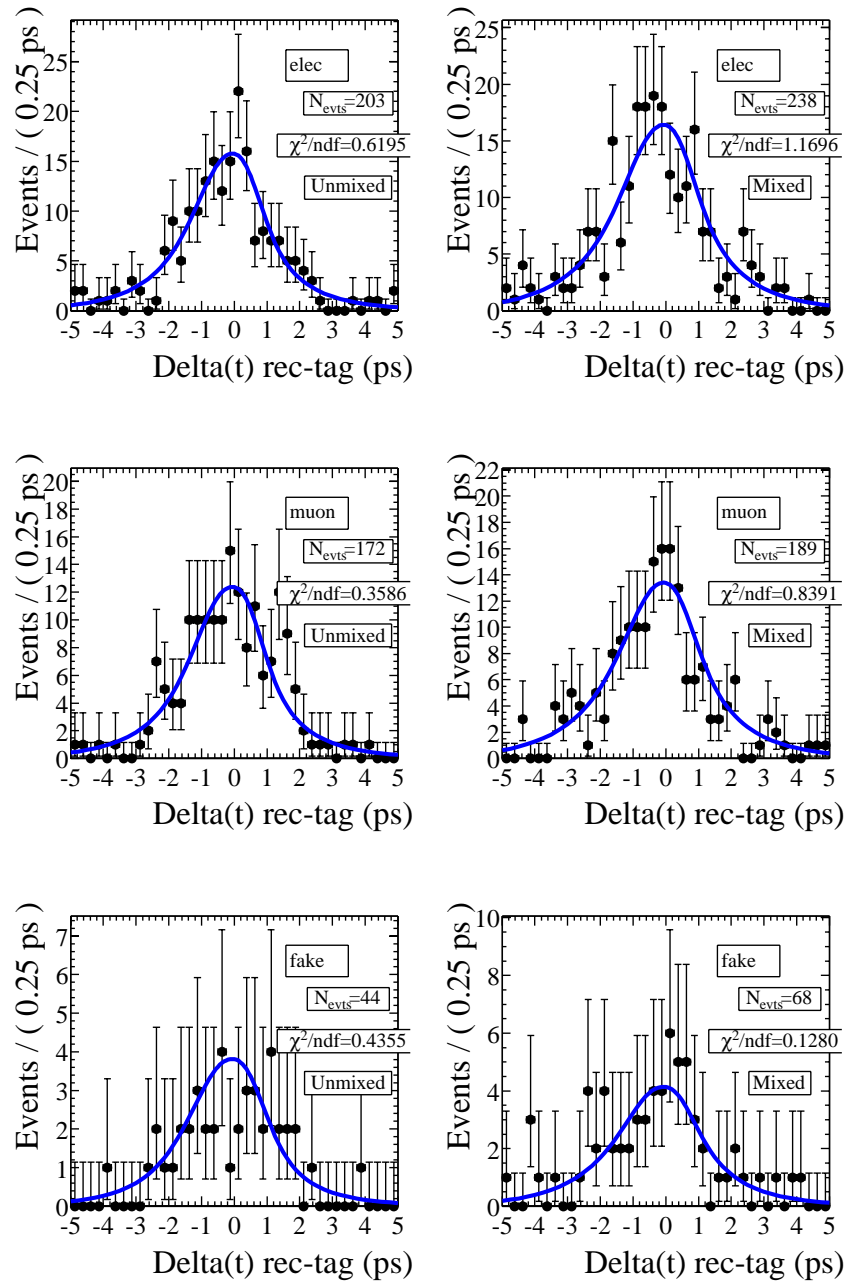
Figure 8.10: Δt projection of the simultaneous fit to the uncorrelated lepton sample separated by leptID and mixing status (mixed, unmixed).
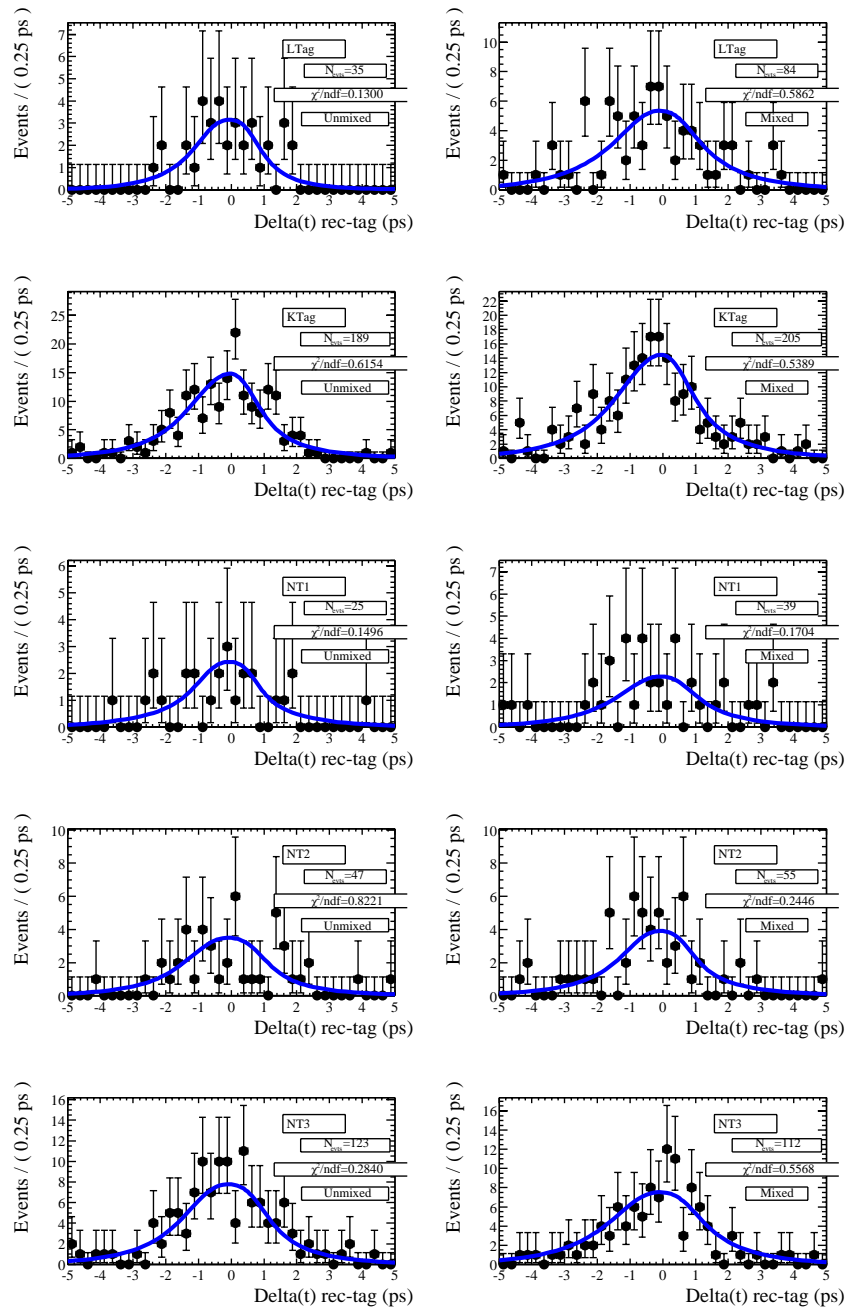
Figure 8.11: Δt projection of the simultaneous fit to the uncorrelated lepton sample separated by tagCat and mixing status (mixed, unmixed).

8.5 Continuum Δt fit

Continuum events are $c\bar{c}$ events where one charm quark produces a D^* while the other charm quark decays to produce a high momentum lepton candidate (could be real or fake lepton). The topology of these events is such that the lepton, D^* and all other decay daughters have a common origin, the decay vertex of the $c\bar{c}$. This implies that these events can very easily be modeled as just having a prompt decay time structure.
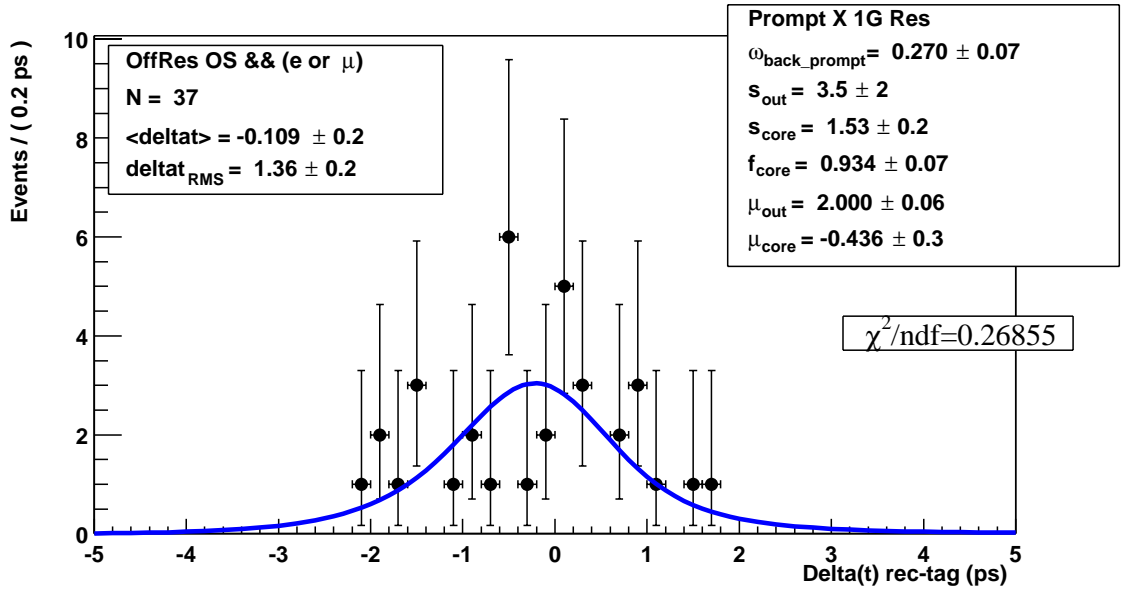
Figure 8.12: The full continuum fit projected into the signal region.

The continuum event time structure will be measured with the off-resonance data sample. This sample is very small compared to the on-resonance sample, it contains less than 2000 events total and only 37 events pass signal region cuts. The continuum events are described by the following PDF:

$$G_{cont}(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,cont}) = G_{prompt} \otimes R_{G+G}(\Delta t, \sigma_{\Delta t}; \vec{r}^{\,cont}) \qquad (8.2)$$

The resolution model has five parameters while the prompt time structure PDF has one parameters (ω_{prompt}). These parameters are broken down by the categories shown in table 8.9. The parameter break downs are limited by the small number of events in the sample.

Parameter	Group	# pars
μ_{core}	OS/SS	2
σ_{core}		1
η_{prompt}	$\text{Kaon} \oplus (L_{or} NT1_{or} NT2_{or} NT3)$	2
f_{outlier}		1
$\text{bias}_{\text{outlier}}$		1
$\text{scale}_{\text{outlier}}$		1

Table 8.9: The final parameter breakdown by categories.

Figures 8.13 and 8.14 shows the results of the fully fitted sample projected into different slices of tagCatGrp and LeptID. The χ^2 values in each plot are either reasonable or extremely low due to low statistics. None of the continuum fit parameters are broken down by signal region or LeptID, so the fact that projections by these regions give reasonable results is reassuring - but mostly a result of the low statistics.

The fitted parameters will have fairly large uncertainties, but this really is no concern, since the measured values will only serve as a starting point for the final fit. The final fit will fit the signal + control samples simultaneously and all Δt parameter will be floating.

8.6   Charged B backgrounds

Charged B backgrounds come primarily from $B^+ \rightarrow D^* l \nu$X. The fraction of charged B background due to these decays in the signal region is approximately 4%, this estimate comes from figure 4.25. The number of charged B's passing the event selection for ≈ 10

fb^{-1} is 510, the total number of events passing event selection is 13007. Charged B's do not mix, thus Δt and tagging information can be used to discriminate against charged B peaking background events. The same resolution model is used for the signal model and charged B events since the decay dynamics between signal and charged B background are very similar. The total PDF is,

$$
\begin{aligned}
G(\Delta t, \sigma_{\Delta t}, g; \vec{q}^{\,sig}) \;=\; & (\frac{1}{4\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}} (1 + g(1 - 2(\omega_{B^0} \pm \Delta\omega_{B^0}))) \cos(\Delta m \Delta t)) + \\
& \frac{1}{4\tau_{B^+}} e^{-|\Delta t|/\tau_{B^+}} (1 + g(1 - 2(\omega_{B^+} \pm \Delta\omega_{B^+})))) \otimes R_{GExp+G} \quad (8.3)
\end{aligned}
$$
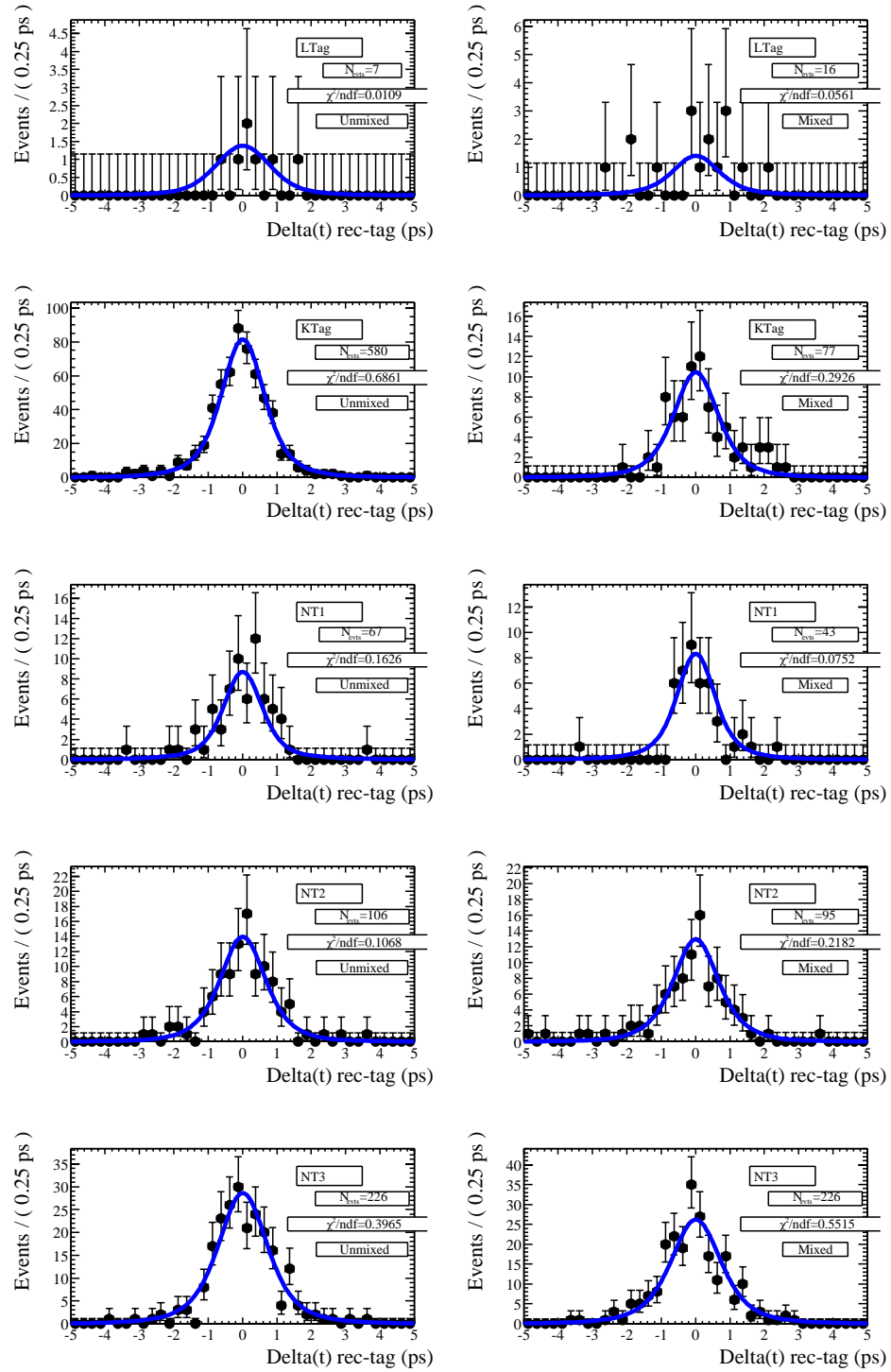
Figure 8.13: The continuum data and PDF projected onto (tagCatGrp⊗mixed/unmixed).
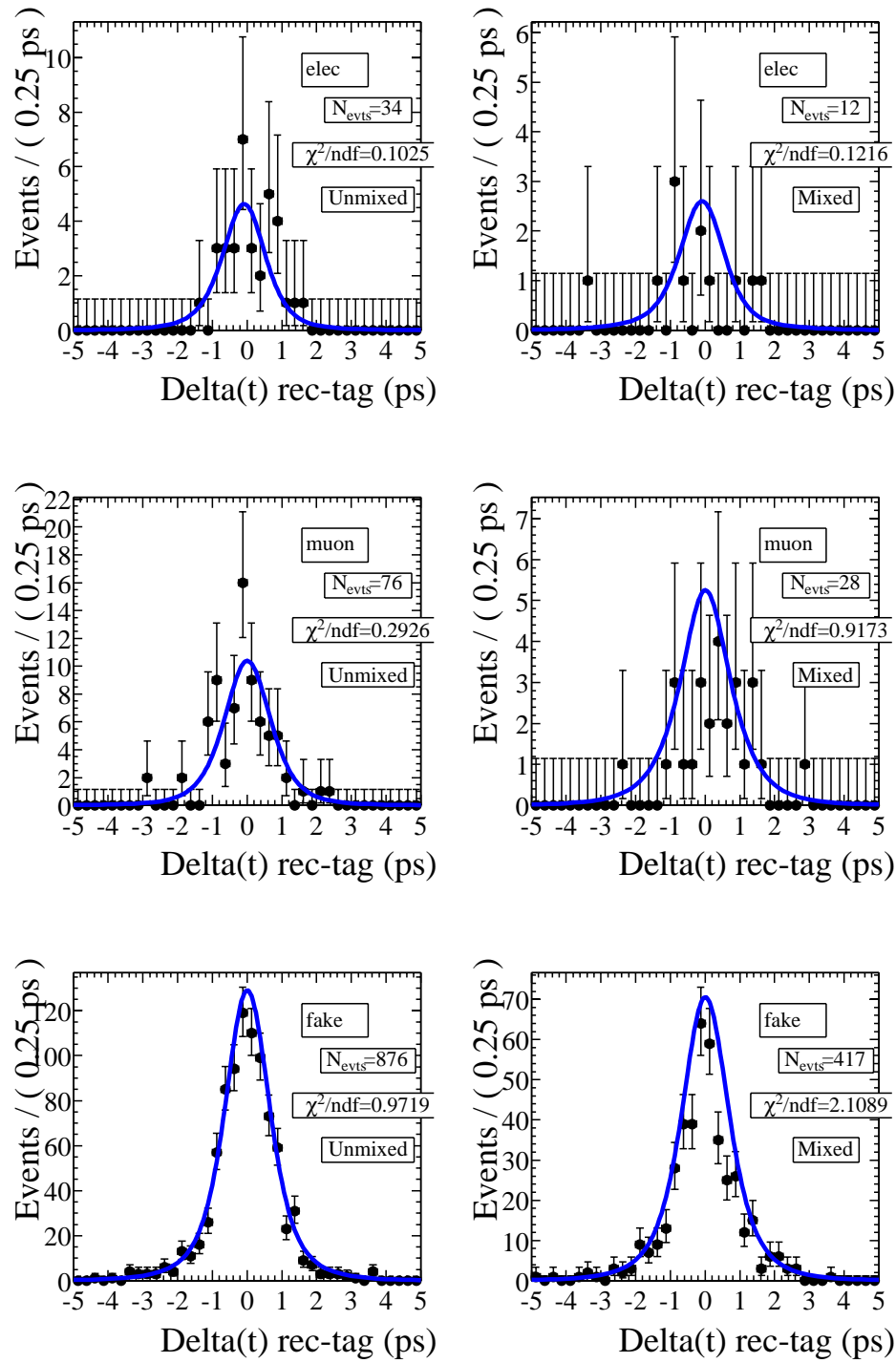
Figure 8.14: The continuum data and PDF projected onto (leptID⊗mixed/unmixed).

Since the ratio of charged B to neutral B lifetime is close to one and only 4% of the events are charged B, there is not sufficient sensitivity to distinguish them. Therefore the ratio is fixed to the PDG value in the fit. The ratio will be varied by the PDG errors to estimate a systematic uncertainty. Either ω_{B^+} or $\omega_{B^+}/\omega_{B^0}$ needs to be fixed because there are only two independent parameters among ω_{B^+}, ω_{B^0} and charged B fraction [3]. The reason is that given that the two lifetimes are so close, the time evolution of mixed and unmixed events can be expressed approximately as,

$$U \quad \propto \quad (1 + f_{B^+}D_{B^+})/2 + ((1 - f_{B^+})D_{B^0}\cos(\Delta m \Delta t))/2$$

$$M \quad \propto \quad (1 - f_{B^+}D_{B^+})/2 - ((1 - f_{B^+})D_{B^0}\cos(\Delta m \Delta t))/2\,,$$

where D is dilution $1 - 2\omega$. Only two quantities, $f_{B^+}D_{B^+}$ and $(1 - f_{B^+})D_{B^0})$ are measurable. Therefore in our fit, the mistag ratio and the lifetime ratio will be fixed.

This model can be used to fit a sample with neutral and charged B mixture from generic Monte Carlo sample, in which 17290 events are real signal and 952 are charge B events, and no other background. The charged B fraction in this mixture is 5.2 %. The lifetime ratio τ_{B^+}/τ_{B^0} and mistag ratio $\omega_{B^+}/\omega_{B^0}$ are set to be 1.069 and 0.97, respectively. The fit result of Δm_d, τ_{B^0} and charged B fraction is shown in Table 8.10. The mistag fraction is based on results from Monte Carlo truth counting of the B^0 and charged B samples. The lifetime fraction is based on the PDG2000 value.

8.7   Full Monte Carlo Fit

This section describes the full fit to generic Monte Carlo data. The total number of floating parameters is 72: 22 in signal model, 24 in combinatoric model and 26 in peaking background model. The ratio of B^+ to B^0 lifetimes is fixed at their generated

Signal Δt fits to (truth-matched) Signal & Generic Monte Carlo				
parameter	Sig MC(*)	Gen B^0	Gen B^0, B^+	Gen B^0,B^+
Δm_d (ps^{-1})	0.4715 ± 0.0043	0.4640 ± 0.0101	0.4630 ± 0.0106	0.4617 ± 0.0132
τ_{B^0} (ps)	1.5523 ± 0.0092	1.5749 ± 0.0227	1.5666 ± 0.0221	1.5656 ± 0.0227
f_{B^+}	0	0	0.0521	0.047 ± 0.030

Table 8.10: Fitted results for Δm_d, τ_{B^0}, and f_{B^+}. The third column of values represents a fit to truth-matched Generic MC events where f_{B^+} is fixed to its true value; in the fourth column this parameter is floated. In all the fits, resolution function and mistag fractions are floating, except for the two signal outlier model shape parameters. (*) Note that these signal MC fit results differ slightly from ones quoted elsewhere in this document; these results are the ones that are most relevant to this table since they incorporate the final fixed values for both of the outlier shape parameters, etc.

value, $r = 1.069$, and the ratios of B^+ to B^0 mistag fractions are fixed at the values from

MC truth counting.

The generic Monte Carlo fit is almost exactly the same as the signal Monte Carlo

fits performed in previous sections. The only significant change is that a new set of fixed

outlier parameters need to be picked. A full scan of the b^{out} and s^{out} parameters is per-

formed to determine reasonable vaules for the outlier parameters.

The results from the likelihood surface scan for Generic MC in the $b^{\mathrm{out}} - s^{\mathrm{out}}$

space is shown in figs 8.15- 8.17.

The fixed values for the Generic MC MasterModel were picked with two impor-

tant factors in mind:

- The signal MC fits using just the signal Δt model preferred outlier shape parame-

  ters ($b^{\mathrm{out}} = -1.68 \pm 0.99$, $s^{\mathrm{out}} = 5.45 \pm 0.86$) ps.

- The weight of the (Δm_d,τ_{B^0}) distribution due to variations in the outlier parameters

  is centered near ($b^{\mathrm{out}}, s^{\mathrm{out}}$) = $(-2, 6)$ and $(0, 8)$ ps. (See Fig. 8.17.)

For the generic Monte Carlo sample, the signal outlier shape parameters are picked to be

($b^{\mathrm{out}}, s^{\mathrm{out}}$) = $(-2, 6)$.

Full Fit to Generic MC, Signal Model Parameters					
parameter	value		value		value
$\Delta m_d\ (ps^{-1})$	0.486 ± 0.015	f_{B^+}	0.095 ± 0.028	s^1	0.981 ± 0.064
τ (ps)	1.580 ± 0.024	ω_{LT}	0.039 ± 0.011	κ	1.19 ± 0.19
-	-	ω_{KT}^{offset}	0.015 ± 0.027	$f^1{}_{LT}$	0.884 ± 0.055
-	-	m_{KT}	0.200 ± 0.039	f^1_{KT}	0.679 ± 0.060
-	-	ω_{N1}	0.180 ± 0.018	$f^1{}_{N1}$	0.670 ± 0.092
-	-	ω_{N2}	0.352 ± 0.015	$f^1{}_{N2}$	0.708 ± 0.073
-	-	ω_{N3}	0.445 ± 0.011	$f^1{}_{N3}$	0.743 ± 0.054
-	-	$\Delta\omega_{LT}$	-0.024 ± 0.018	f^{out}	-0.002 ± 0.002
-	-	$\Delta\omega_{KT}$	-0.026 ± 0.014	b^{out}	-2.00
-	-	$\Delta\omega_{N1}$	0.012 ± 0.029	s^{out}	6.00
-	-	$\Delta\omega_{N2}$	-0.004 ± 0.024	-	-
-	-	$\Delta\omega_{N3}$	-0.096 ± 0.018	-	-

Table 8.11: Full fit to generic MC — signal model parameters. (Recall that s^{out} and b^{out} for the signal resolution model are fixed to constant values.)

Fit parameters for signal and background models are shown in tables 8.11- 8.14. Several things can be observed:

- The observed outlier fraction is low by, perhaps, about 1σ.

- The measured Δm_d value is 1σ high, the measured τ_{B^0} value is 1.4σ too high, each with respect to the true values from the generator level. They are 1.5σ and 0.58σ too high (respectively) compared to the values from the truth-matched sample of Generic B^0 and B^+.

- The outlier fraction is still consistent with zero, as originally observed in the fits to only only truth-matched candidates from Generic MC.

- The charged B fraction, f_{B^+}, is significantly higher, by about 1.6σ with respect to the true value, or 1.65σ with respect to the result from the truth-matched fit.

- The observed slope for the linear relation between m_{KT} and $\sigma_{\Delta t}$ is significantly higher, by 2.1σ. Additionally, all the mistag rates seems perturbed by at least 1σ,

Full Fit to Generic MC, Comparing Signal Model Pars

$\delta(par)$	value	$\delta(par)$	value	$\delta(par)$	value
$\Delta m_d\ (ps^{-1})$	0.014 ± 0.015	f_{B+}	0.043 ± 0.028	s^1	0.026 ± 0.064
τ (ps)	0.032 ± 0.024	ω_{LT}	-0.015 ± 0.011	κ	0.041 ± 0.19
-	-	$\omega_{\mathrm{KT}}^{\mathrm{offset}}$	-0.062 ± 0.027	$f^1{}_{\mathrm{LT}}$	0.040 ± 0.055
-	-	m_{KT}	0.081 ± 0.039	f^1_{KT}	-0.023 ± 0.060
-	-	ω_{N1}	0.041 ± 0.018	$f^1{}_{\mathrm{N1}}$	-0.067 ± 0.092
-	-	ω_{N2}	0.041 ± 0.015	$f^1{}_{\mathrm{N2}}$	-0.057 ± 0.073
-	-	ω_{N3}	-0.029 ± 0.011	$f^1{}_{\mathrm{N3}}$	0.011 ± 0.054
-	-	$\Delta\omega_{\mathrm{LT}}$	-0.015 ± 0.018	f^{out}	0.002 ± 0.002
-	-	$\Delta\omega_{\mathrm{KT}}$	-0.020 ± 0.014	b^{out}	0.32
-	-	$\Delta\omega_{\mathrm{N1}}$	0.041 ± 0.029	s^{out}	0.55
-	-	$\Delta\omega_{\mathrm{N2}}$	0.041 ± 0.024	-	-
-	-	$\Delta\omega_{\mathrm{N3}}$	-0.029 ± 0.018	-	-

Table 8.12: Full fit to generic MC — comparison of signal model results to generator truth (for Δm_d, τ_{B^0}, and f_{B+}) and truth-matched generic MC results for everything else. (Recall that s^{out} and b^{out} for the signal resolution model are fixed to constant values.) Errors listed are the error on the Gen MC fit result, the appropriate σ for measuring degree of difference.

and this is an underestimate because the truth-match signal events are a subset of the full Generic MC data sample.

The large value of f_{B+} is surprising, in particular since the result from the full MasterModel fit should be highly correlated with the result from the fit to only the truth-matched signal events. One problem which could cause this is that the Generic MC sample has an unusually small population of outliers in the large Δt tails which degrades the ability to fit for their overall fraction. This inability, in turn, weakens the signal resolution model which convolutes the charged and neutral B time structures together.

The correlation coefficients in the final MasterModel Generic MC result shown in Table 8.16. An upward fluctuation in f_{B+} by 1.6σ would induce a downward flucctuation in f^{out} of about 0.5σ, and cause a positive change in Δm_d of about 0.8σ, The low f^{out} could induce an increase in τ_{B^0} of about 0.4σ. These shifts are not quite sufficient to explain the parameters observed in the full fit result, however. If, however, f_{B+} is fixed

Full Fit to Generic MC, Combinatoric Δt Model Parameters			
parameter	value	parameter	value
Δm_{comb}	0.394 ± 0.018	$b^1{}_{\text{LNT12}}$	-0.060 ± 0.022
τ_{comb}	1.275 ± 0.020	$b^1{}_{\text{KNT3}}$	-0.159 ± 0.012
$\omega^{\text{osc}}_{\text{Fake;LTag}}$	0.367 ± 0.023	$f^{\text{osc}}_{\text{Off}}$	0.150 ± 0.053
$\omega^{\text{osc}}_{\text{Fake;KTag}}$	0.194 ± 0.013	$f^{\text{osc}}_{\text{SS}}$	0.469 ± 0.025
$\omega^{\text{osc}}_{\text{Fake;NT1}}$	0.381 ± 0.027	$f^{\text{osc}}_{\text{NoLept}}$	0.686 ± 0.018
$\omega^{\text{osc}}_{\text{Fake;NT2}}$	0.444 ± 0.017	$f^{\text{osc}}_{\text{OneLept}}$	0.900 ± 0.018
$\omega^{\text{osc}}_{\text{Fake;NT3}}$	0.533 ± 0.014	$f^{\text{osc}}_{\text{TwoLept}}$	1.000
$\omega^{\text{osc}}_{\text{Lept;LTag}}$	0.215 ± 0.023	$s^1{}_{\text{OffSS}}$	1.242 ± 0.030
$\omega^{\text{osc}}_{\text{Lept;KTag}}$	0.258 ± 0.010	$s^1{}_{\text{On;OS}}$	1.129 ± 0.025
$\omega^{\text{osc}}_{\text{Lept;NT1}}$	0.269 ± 0.024	f^1	0.9837 ± 0.0037
$\omega^{\text{osc}}_{\text{Lept;NT2}}$	0.426 ± 0.016	b^{out}	-1.57 ± 0.42
$\omega^{\text{osc}}_{\text{Lept;NT3}}$	0.494 ± 0.011	s^{out}	6.80 ± 0.53
ω^{pmt}	0.427 ± 0.012	-	-

Table 8.13: Full fit to generic MC — combintoric Δt model parameters.

to lower values, performance can be recovered, as illustrated in Table 8.15.

Full Fit to Generic MC, Peaking BG Model Parameters					
cont. par.	value	fake par.	value	unco. par.	value
ω_{KT}	0.076 ± 0.029	Δm^{fake}	0.405 ± 0.023	τ^{unco}	1.68 ± 0.35
ω_{LNT}	0.503 ± 0.037	τ^{fake}	1.386 ± 0.038	$\omega_{\mathrm{KNT}}^{\mathrm{life}}$	0.67 ± 0.13
$b^1{}_{\mathrm{OS}}$	-0.022 ± 0.104	$\omega_{\mathrm{LT}}^{\mathrm{osc}}$	0.192 ± 0.025	$\omega_{\mathrm{LT}}^{\mathrm{life}}$	0.999
$b^1{}_{\mathrm{SS}}$	-0.092 ± 0.111	$\omega_{\mathrm{KT}}^{\mathrm{osc}}$	0.180 ± 0.020	ω^{pmt}	0.58 ± 0.07
s^1	1.286 ± 0.076	$\omega_{\mathrm{NT1}}^{\mathrm{osc}}$	0.278 ± 0.029	f^{life}	0.37 ± 0.12
f^1	0.963 ± 0.024	$\omega_{\mathrm{NT2}}^{\mathrm{osc}}$	0.381 ± 0.028	b^1	-0.44 ± 0.13
-	-	$\omega_{\mathrm{NT3}}^{\mathrm{osc}}$	0.508 ± 0.024	f^1	0.993
-	-	$\omega_{\mathrm{KNT3}}^{\mathrm{pmt}}$	0.252 ± 0.086	s^1	1.33 ± 0.17
-	-	$\omega_{\mathrm{LNT12}}^{\mathrm{pmt}}$	0.471 ± 0.106	-	-
-	-	f^{osc}	0.881 ± 0.027	-	-
-	-	$b^1{}_{\mathrm{KNT3}}$	-0.216 ± 0.041	-	-
-	-	$b^1{}_{\mathrm{LNT12}}$	-0.043 ± 0.045	-	-
-	-	f^1	0.965 ± 0.010	-	-
-	-	s^1	1.062 ± 0.068	-	-

Table 8.14: Full fit to generic MC — peaking background Δt model parameters. The outlier bias (b^{out}) and scale factor (s^{out}) for these peaking background are shared with combinatoric background parameters.

Full fit results to Generic MC w/fixed f_{B^+}						
f_{B^+}	$\delta(f_{B^+})$	Δm_d	$\delta(\Delta m_d)$	τ_{B^0}	$\delta(\tau_{B^0})$	f^{out}
0.0474	-0.0047	0.4718	-0.0002	1.5718	0.0238	-0.00008
0.0521	0.0000	0.4732	0.0012	1.5728	0.0248	-0.0003
0.0600	0.0079	0.4756	0.0036	1.5744	0.0264	-0.0006
0.0700	0.0179	0.4786	0.0066	1.5764	0.0284	-0.0011
0.0951	0.0430	0.4856	0.0136	1.5806	0.0326	-0.0019

Table 8.15: Selected results from full MasterModel fits to Generic MC with fixed f_{B^+} to several different values. There are strong correlations between these four parameters. The columns labelled $\delta()$ reflect the change with respect to the value from MC truth. Recall that the statistical error on Δm_d from these fits is around 0.015 ps^{-1} and about 0.024 ps for τ_{B^0}.
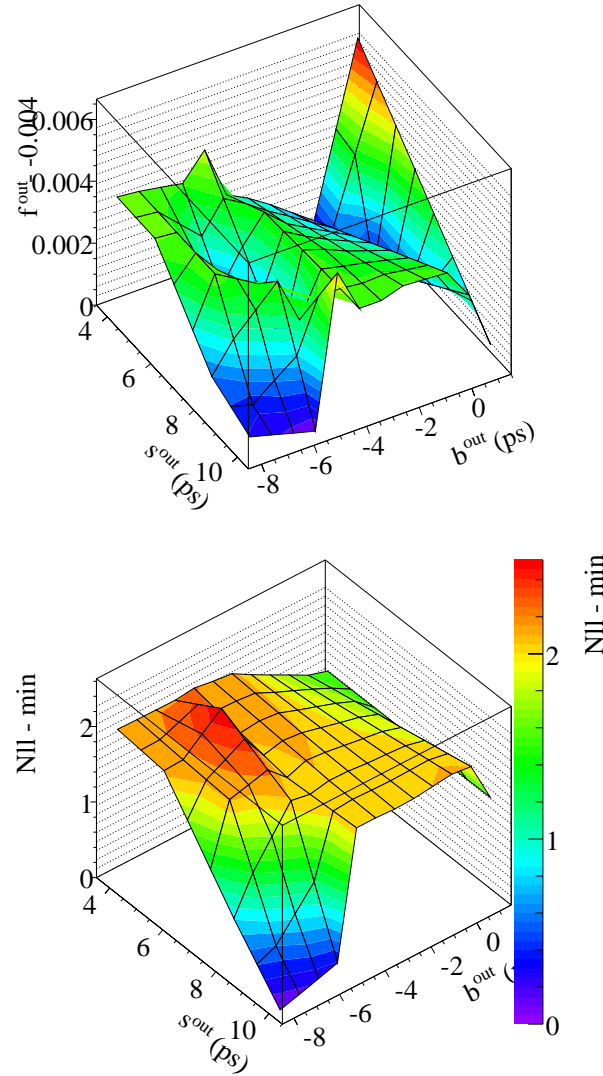
Figure 8.15: The negative-ln-likelihood surface (right) and f^{out} surface (left) in the space of bias and outlier width of the outlier signal resolution model, as fitted in Generic MC. The vertical scale has been offset so that minimum is at 0 (note that the minimum for f^{out} is at -0.004, so the surface has been moved by -0.004 units). The surface is constructed from 33 fits to Generic MC with different fixed values of the signal outlier parameters. This surface is essentially flat until the edges where the extreme values of the model make it unstable, *i.e.* the fitted outlier fraction becomes significantly less than 0. A physical, preferred minimum is not observed.
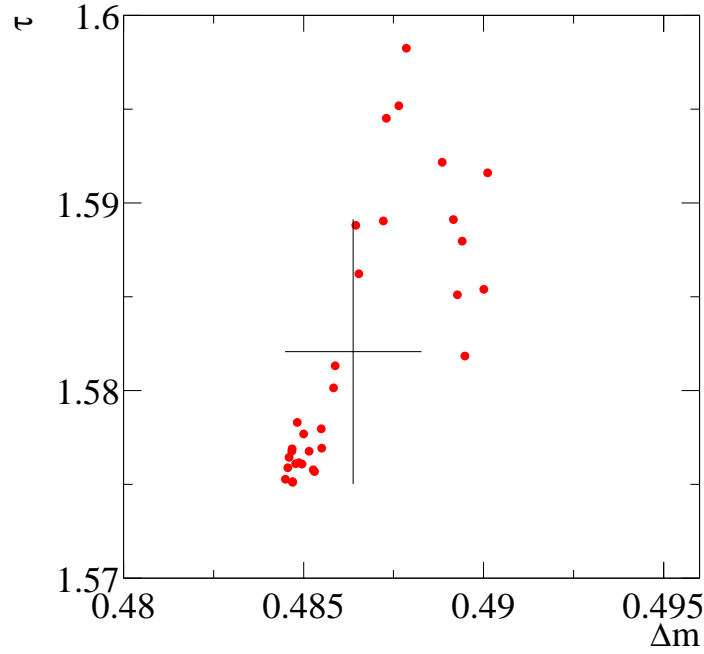
Figure 8.16: Scatter plot of the set of Δm_d, τ_{B^0} values obtained from data for 33 fits with different fixed values of the outlier bias and width in the range 1 to -7 ps for bias and 4 to 10 ps for width. The cross indicates the mean and rms of the Δm_d, τ_{B^0} distribution. Recall that the statistical error from the full fit is 0.023 ps for τ_{B^0} and 0.014 ps^{-1} for Δm_d for the Generic MC fit. Thus the variation in τ_{B^0} and Δm_d due different assumptions about the fixed values of the signal outlier shape is comparatively minor.

Corr. coeff's from full fit to Gen MC			
pars	value	pars	value
$\rho(\Delta m_d, \tau_{B^0})$	-0.162	$\rho(m_{\mathrm{KT}}, \omega_{\mathrm{KT}}^{\mathrm{offset}})$	-0.944
$\rho(f^{\mathrm{out}}, \Delta m_d)$	-0.169	$\rho(f_{B^+}, \Delta m_d)$	0.552
$\rho(f^{\mathrm{out}}, \tau_{B^0})$	-0.402	$\rho(f_{B^+}, \tau_{B^0})$	0.162
$\rho(f^{\mathrm{out}}, m_{\mathrm{KT}})$	-0.018	$\rho(f_{B^+}, m_{\mathrm{KT}})$	0.086
$\rho(f^{\mathrm{out}}, \omega_{\mathrm{KT}}^{\mathrm{offset}})$	0.004	$\rho(f_{B^+}, \omega_{\mathrm{KT}}^{\mathrm{offset}})$	-0.060
$\rho(f^{\mathrm{out}}, f_{B^+})$	-0.333		

Table 8.16: Correlation coefficients from baseline MasterModel fit to Generic MC. (f_{B^+} is correlated with all the mistag rates at only the few % level.)
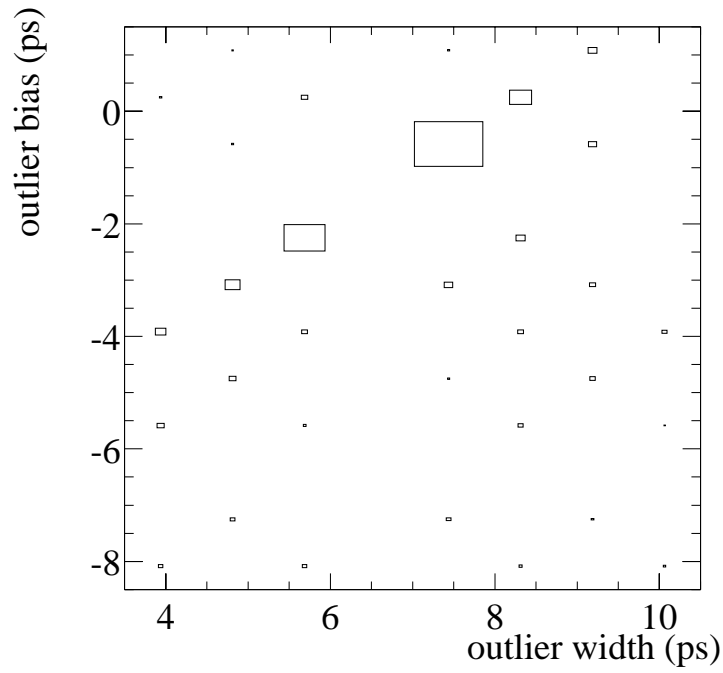
Figure 8.17: Distribution of full MasterModel fit results for fits to Generic MC. The grid represents the space of different assumptions about the two outlier shape parameters, and the weight assigned to each point represents the proximity of the $(\Delta m_d, \tau_{B^0})$ fit results to the mean and rms mentioned in the previous plot. From this distribution, the grid point (-2,6) and (0,8) are preferred as being sets of outlier shape parameters that yield Δm_d, τ_{B^0} fit results most near the center of their distribution.

8.8   Monte Carlo Correction

This section describes how generic Monte Carlo data is used to estimate the systematic difference in fit values due to the full fitting procedure. These estimates will be used as a correction to get an unbiased estimate of the fit values, the uncertainties in the corrections will contribute to the systematic error of the measurement.

High statistics fits using signal Monte Carlo (instead of generic Monte Carlo) have shown that the signal PDF gives an unbiased fit. The strategy for determining the Monte Carlo correction is as follows:

- Fit only Monte Carlo truth match signal events with the signal PDF only.

- Fit the full generic sample with the full fit.

- The difference represent a systematic bias which gives the Monte Carlo correction

The bias is $+0.022$ ps for τ_{B^0} and $+0.020$ ps^{-1} for Δm_d. The systematic error due to the statistical uncertainty on this correction is conservatively taken to be the full statistical error of the fit to signal+background: ± 0.018 ps for τ_{B^0} and ± 0.012 ps^{-1} for Δm_d.

8.9   Toy Study

The fitting procedure is also tested with a Monte Carlo toy study. Toy Monte Carlo data sets with the exact same number of events as the real data set are produced using the same central values as determined from the full data fit. The values are the blinded fit values, the study was set up so that the values would be unblinded internally to generate the toy data sets. Subsequent fits would use the same blinding string to

Fits to Generic Monte Carlo Sample						
Type	τ_{B^0}	*diff*(10^{-3})	Δm_d	*diff*(10^{-3})	change w.r.t.
generator	1.548	-	0.472	-	-
truth fit	1.538 ± 0.011	-10 ± 11	0.469 ± 0.005	-3 ± 5	generator
signal	1.566 ± 0.023	$+28 \pm 20$	0.462 ± 0.013	-7 ± 12	truth fit
full model	1.580 ± 0.024	$+14 \pm 6$	0.486 ± 0.015	$+24 \pm 6$	signal
Fits to Generic + Filtered Monte Carlo Sample						
Type	τ_{B^0}	*diff*(10^{-3})	Δm_d	*diff*(10^{-3})	change w.r.t.
truth fit	1.556 ± 0.009	$+8 \pm 9$	0.470 ± 0.004	-2 ± 4	generator
signal	1.554 ± 0.016	-2 ± 13	0.459 ± 0.011	-12 ± 10	truth fit
full model	1.576 ± 0.018	$+22 \pm 9$	0.479 ± 0.012	$+20 \pm 5$	signal

Table 8.17: Fit results and biases for generic Monte Carlo sample and for generic + filtered Monte Carlo sample. The errors on changes are calcuated as difference in quadrature between two errors.

produce consistent results.

An ensemble of 40 Monte Carlo data sets are generated and fitted. The distributions of the fit values of Δm_d, τ_{B^0} and f_{B^+} are shown in Fig. 8.18. The mean and RMSs of Δm_d and τ_{B^0} should be compared with the central value and statistical error on the (blind) fit values from data:

$$\Delta m_d = 0.527 \pm 0.018 \; ps^{-1} \; \text{(blind)} \tag{8.4}$$

$$\tau_{B^0} = 1.539 \pm 0.023 \; \text{ps (blind)} \tag{8.5}$$

The central values of Δm_d, τ_{B^0} and f_{B^+} from the fit to data (used to generate the toy MC samples) are shown as vertical dashed lines in Fig. 8.18. The means are very consistent, but the RMS's are almost 20% larger then the statistical errors from the fit. This inflation might be due to a problem concerning refitting the δm variable which isn't correctly done in the toy study.

In order to make sure the statistical errors of the fit to data are correct, a log-likelihood scan is performed and a small correction on the statistical errors is made. The correction is much smaller than 20%.
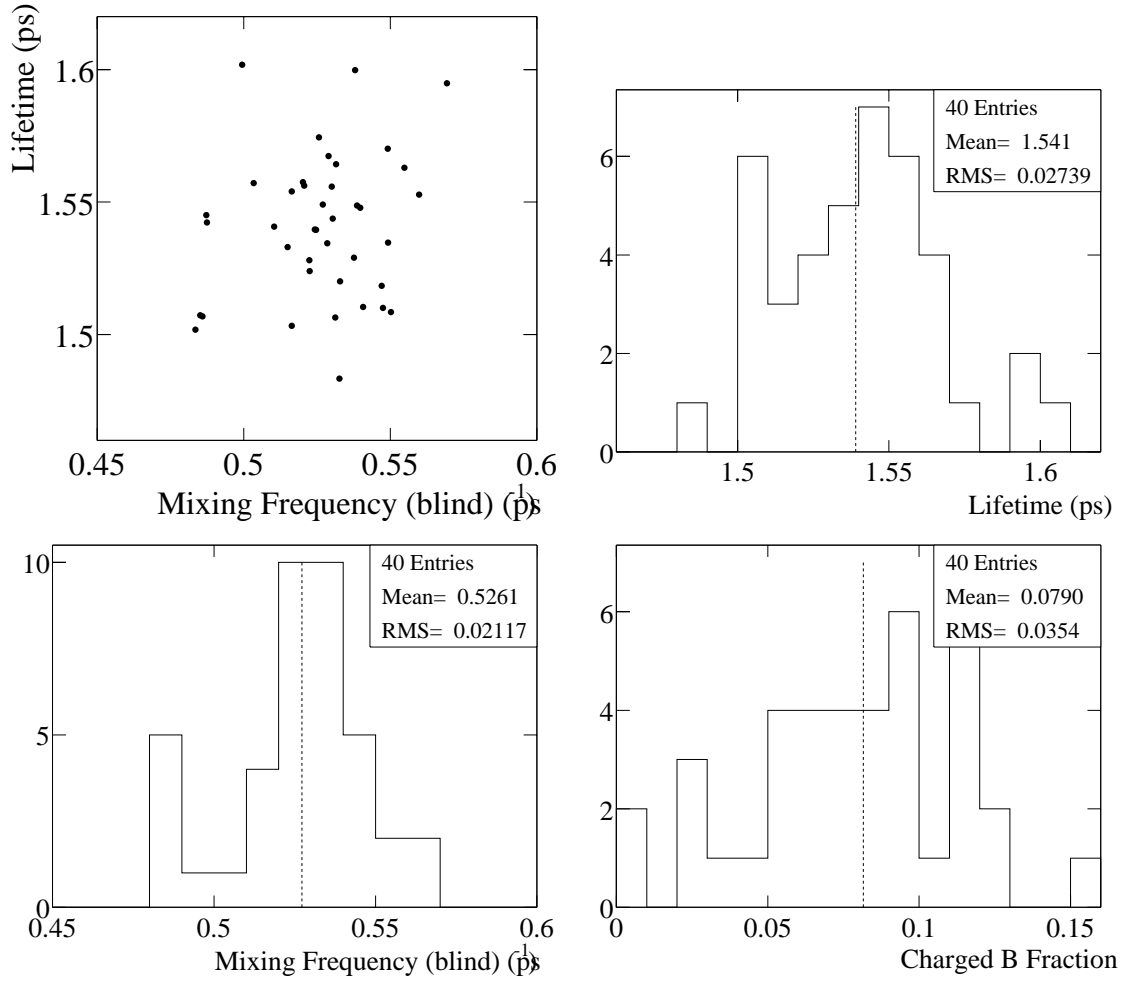
Figure 8.18: The distributions of τ_{B^0}, Δm_d and f_{B^+} from 40 toy Monte Carlo sample fits. The values from the default fit to data, which were used to generate the toy MC samples, are indicated by vertical dashed lines.

The distribution of minimized negative log-likelihood (minNll) is shown in Fig. 8.19.

The number of toy MC samples that result in a smaller (better) value of minNll is 24 (out

of 40). The consistency between the model and the data is very reasonable.
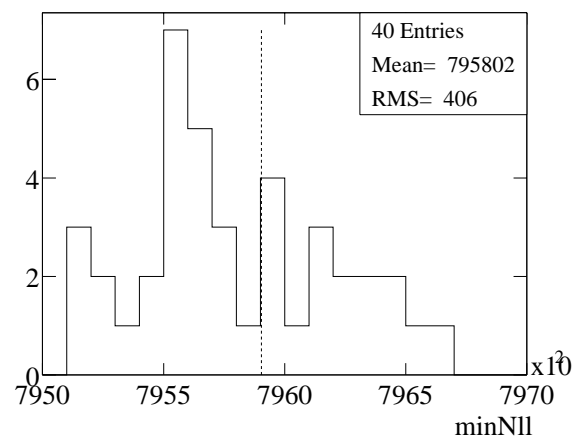
Figure 8.19: The distribution of minimized negative log-likelihood (minNll) from 40 toy Monte Carlo sample fits. The value from default fit to data is indicated by the vertical dashed lines. The number of fits that result in a smaller (better) value of minNll is 24.

Chapter 9

Consistency Checks

This section describes the cross checks which have been performed to confirm the validity of the fit results and check for any possible problems in the data set. Both Monte Carlo and real data events are used to perform the checks. Many checks involve looking at various sub-samples to confirm that fit results from the sub-samples are consistent with the full fit.

The fits in this section will generally be compared to a reference fit and only the differences are shown in tables. Monte Carlo fits will be compared to the full signal Monte Carlo fit shown in the last column of table 9.1. Results involving real data events will be compared to the full data fit shown in Tables 11.2, 11.3 and 11.4 in Sec. 11 using the Master Model. If the differences between fit and reference parameters are essentially zero for a group of parameters, this will be noted with a single line entry in the table.

9.1 Fit to MC truth

Fits using the Monte Carlo truth values of Δt and Δz were performed in section 8.1. The results are summarized in Table 9.1. For the MC truth Δt fits, the value of

Sample	generator value	MCtruth Δt fit	MCtruth Δz fit	signal model fit
Δm_d (ps^{-1})	0.472	0.4708$\pm$0.0013	0.4723$\pm$0.0007	0.4715$\pm$0.0043
τ_{B^0} (ps)	1.548	1.5446$\pm$0.0051	1.5494$\pm$0.0051	1.5523$\pm$0.0091
ω_{LT}	-	0.063	0.063	0.056$\pm$0.003
$\omega_{\mathrm{KT}}^{\mathrm{offset}}$	-	0.157	0.157	0.067$\pm$0.008
m_{KT}	-	-	-	0.127$\pm$0.012
ω_{N1}	-	0.192	0.192	0.184$\pm$0.006
ω_{N2}	-	0.338	0.338	0.339$\pm$0.005
ω_{N3}	-	0.465	0.465	0.462$\pm$0.004

Table 9.1: Fitting for τ_{B^0}, mistag rate and Δm_d using signal MC candidates. The MC truth fits use truth information from all correctly reconstructed events in the signal MC sample. The standard fit uses reconstruction information from all correctly reconstructed signal MC events and fixes the outlier bias and width. The MC truth mistag rates are calculated by using the reconstructed and truth mixing status for all events.

τ_{B^0} deviates from the generator value of 1.548 ps by -0.0030$\pm$0.0051 ps (about -0.6σ) while the value of Δm_d deviates from the generator value of 0.472 ps^{-1} by -0.0014$\pm$0.0012 ps^{-1} (about -1.2σ). For the MC truth Δz fits, the value of τ_{B^0} deviates from the generator value by 0.0014$\pm$0.0051 ps (about 0.3σ) while the value of Δm_d deviates from the generator value by 0.0004$\pm$0.0012 ps^{-1} (about 0.3σ). A plot of τ_{B^0} fit is shown in Fig. 5.16 using MC truth Δt and Δz information.

9.2 True vs Reconstructed Flavor Tagging

One cross check performed is comparing the mistag rates determined from a fit vs the mistag rates determined by counting the fraction of correctly tagged events using MC truth information. The value of χ_d can also be determined with a simple counting technique, this was done in section 4.10. Table 9.1 shows the measured and counted mistag rates for each tagging category.

The MC truth fit uses both true Δt and true tagging information. The last column is the "standard" fit which used both reconstructed Δt and tagging information.

Only the last fit uses a resolution model, since the other fits have perfect Δt information. Also, the kaon mistag dependence on $\sigma_{\Delta t}$ is only valid for the last fit, the other fits don't have per-event-errors.

9.3 Δm_d **from** Δt**-Shape or Flavor Information Only**

The Δm_d measurement uses both Δt and flavor tagging information simultaneously to determine Δm_d and τ_{B^0}. One consistency check involves finding the value of Δm_d only using Δt-shape information and finding Δm_d using only flavor tagging information (the number of mixed and unmixed events gives a "time-integrated" mixing measurement). The full signal Monte Carlo sample is used because it has the largest sample size which will yield the highest precision for comparing results.

To estimate Δm_d using only Δt-shape information, the mixing information is eliminated as a dependent variable in the fit. To get the most comparable answer with the the flavor-only result, all parameters in the Δt-only fit are constant except for Δm_d. For signal Monte Carlo, $\Delta m_d^{(shape-only)} = 0.47627 \pm 0.00487$ ps^{-1}, which can be compared with the full fit result of $\Delta m_d = 0.47257 \pm 0.00357$ (when only floating Δm_d).

The value of Δm_d can be calculated with just flavor information by using χ_d which involves counting mixed and unmixed events. This yields $\chi_d = 0.174 \rightarrow \Delta m_d = 0.4705$ using the measured τ_{B^0} and in the absence of mistag rates. An alternative technique is to return the full signal fit, but fix the resolution function to a very, very wide Gaussian (width > 18 ps) so that all Δt information is significantly smeared, and re-fit for Δm_d only. In this case, $\Delta m_d^{(flavor-only)} = 0.47361 \pm 0.00533$ ps^{-1}.

The two results are consistent with the full fit and have similar statistical error. This shows that the shape information and flavor tagging information are both important

in the Δm_d measurement and have similar size errors.

9.4 Subsamples by Tagging Category

The full fit for signal Monte Carlo was performed for each tagging category, a common outlier and core scale factor are used. The mistag rate, core bias, τ_{B^0} and Δm_d are fit inside each tagging category. The mistag rate is made up of two parts, a constant offset, ω^{offset}, and a scale factor which multiplies the per-event-error, m. Thus $\omega = \omega^{offset} + m \cdot \sigma_{\Delta t}$. The scale factor is non-zero only for the kaon tagging category. The results are consistent across the five tagging categories, as shown in Table 9.2 for signal MC and Table 9.3 for the full fit to Data. The values of τ_{B^0} and Δm_d for each tagging category are consistent with the baseline results.

Signal Δt fit (GExp+G) to signal MC					
par.	LTag	KTag	NT1	NT2	NT3
$\delta(\Delta m_d)(\times 10^3)$	-4 ± 7	0 ± 6	11 ± 13	0 ± 19	12 ± 49
$\delta(\tau_{B^0})(\times 10^3)$	-38 ± 16	0 ± 12	7 ± 21	0 ± 17	29 ± 14
$\delta(\Delta\omega)(\times 10^3)$	0 ± 5	0 ± 6	0 ± 10	0 ± 8	0 ± 6
$\delta(\omega^{\text{offset}})(\times 10^3)$	1 ± 4	0 ± 9	-2 ± 6	0 ± 6	0 ± 4
$\delta(m_{\text{KT}})(\times 10^3)$		0 ± 12			
$\delta(f^1)(\times 10^3)$	-13 ± 25	-8 ± 24	-6 ± 31	-7 ± 27	3 ± 22
$\delta(\kappa)(\times 10^3)$			-24 ± 72		
$\delta(s^1)(\times 10^3)$			-6 ± 2		
$\delta(f^{\text{out}})(\times 10^3)$			0 ± 1		

Table 9.2: The table shows the difference in the fit results with respect to the baseline result which uses the combined statistical power of all tagging categories. The fit is done with the GExp+G resolution model. The units for Δm_d and τ_{B^0} are ps^{-1} and ps. The lower section of the table shows parameters which are split on tagging category and are common for the whole sample.

par.	Full Δt fit (GExp+G) to Data				
	LTag	KTag	NT1	NT2	NT3
$\delta(\Delta m_d)(\times 10^3)$	-47 ± 26	16 ± 22	36 ± 45	12 ± 88	9 ± 183
$\delta(\tau_{B^0})(\times 10^3)$	11 ± 48	7 ± 32	-3 ± 59	23 ± 50	-66 ± 36
$\delta(\Delta\omega)(\times 10^3)$	-3 ± 4	0 ± 3	-2 ± 5	1 ± 8	0 ± 3
$\delta(\omega^{\text{offset}})(\times 10^3)$	7 ± 4	-5 ± 3	-8 ± 7	-1 ± 6	-1 ± 3
$\delta(m_{\text{KT}})(\times 10^3)$		0 ± 5			
$\delta(f^1)(\times 10^3)$	125 ± 77	113 ± 73	100 ± 97	100 ± 76	66 ± 64
$\delta(\kappa)(\times 10^3)$			-27 ± 79		
$\delta(s^1)(\times 10^3)$			9 ± 12		
$\delta(f^{\text{out}})(\times 10^3)$			-2 ± 1		

Table 9.3: Master Model fits to the full Dataset using separate Δm_d and τ_{B^0} parameters for each tagging category. The table shows the difference in the fit results with respect to the baseline result. The fit is done with the GExp+G resolution model. The units for Δm_d and τ_{B^0} are ps^{-1} and ps. The errors shown are the errors from the fit. The lower section of the table shows parameters which are already split on tagging category and are common for the whole sample.

9.5 Subsamples by B^0 Flavor

The sample was split into two subsamples, according to the reconstructed flavor of the signal $B^0, \overline{B^0}$ and by tag side $B^0, \overline{B^0}$ as a consistency check. The results are shown in Table 9.4 for signal MC, and in Table 9.5 for Data. The differences in mistag rates for the Rec B^0, Rec $\overline{B^0}$ categories are consistent with the Δ Dilutions measured in the standard fit. The differences in the Tag B^0, Tag $\overline{B^0}$ samples come from the fact that most events are unmixed, thus there is a strong correlation between an event having a Tag B^0 and a Rec $\overline{B^0}$. The correlation isn't perfect, so the difference in mistag rates isn't as large. The values of the two physics parameters appear to be consistent across the subsamples.

	Signal Δt fits (GExp+G) to signal MC subsamples			
	Reconstruction side		Tag side	
	Rec B^0	Rec $\overline{B}^0$	Tag B^0	Tag $\overline{B}^0$
$\delta(\Delta m_d)$ (ps^{-1})	-0.002 ± 0.005	0.002 ± 0.004	0.007 ± 0.004	-0.006 ± 0.005
$\delta(\tau_{B^0})$ (ps)	-0.009 ± 0.010	0.010 ± 0.009	0.008 ± 0.008	-0.009 ± 0.009
$\delta(\omega_{\mathrm{LT}})$	0.001 ± 0.004	-0.002 ± 0.003	0.002 ± 0.004	-0.003 ± 0.003
$\delta(\omega_{\mathrm{KT}}^{\mathrm{offset}})$	-0.012 ± 0.009	0.012 ± 0.008	-0.007 ± 0.009	0.007 ± 0.008
$\delta(m_{\mathrm{KT}})$	0.019 ± 0.013	-0.019 ± 0.012	0.013 ± 0.013	-0.013 ± 0.012
$\delta(\omega_{\mathrm{N1}})$	0.005 ± 0.006	-0.006 ± 0.006	0.013 ± 0.006	-0.014 ± 0.006
$\delta(\omega_{\mathrm{N2}})$	-0.002 ± 0.005	0.002 ± 0.005	0.007 ± 0.006	-0.006 ± 0.005
$\delta(\omega_{\mathrm{N3}})$	0.002 ± 0.004	-0.002 ± 0.004	0.012 ± 0.004	-0.010 ± 0.004
	All Resolution Parameters Consistent			

Table 9.4: Signal Δt fits on subsamples of correctly reconstructed signal MC, separated by tagged or reconstructed flavor of the B^0, reconstructed $B^0,\overline{B}^0$ or tag side $B^0,\overline{B}^0$. The difference in central values is shown with respect to the nominal fit result, and the errors shown are the uncorrelated errors. The fits use the standard GExp+G resolution function. The results are consistent between the flavor subsamples.

	Full Δt fits (GExp+G) to Data			
	Reconstruction side		Tag side	
	RecB^0	Rec$\overline{B}^0$	TagB^0	Tag$\overline{B}^0$
$\delta(\Delta m_d)$ (ps^{-1})	0.026 ± 0.018	-0.041 ± 0.018	-0.045 ± 0.010	0.001 ± 0.015
$\delta(\tau_{B^0})$ (ps)	-0.047 ± 0.025	0.041 ± 0.023	-0.015 ± 0.023	0.019 ± 0.025
f^{out}	0.005 ± 0.002	0.001 ± 0.019	0.003 ± 0.002	0.001 ± 0.003
$\delta(\omega_{\mathrm{LT}})$	-0.011 ± 0.017	0.024 ± 0.012	-0.010 ± 0.011	0.028 ± 0.013
$\delta(\omega_{\mathrm{KT}}^{\mathrm{offset}})$	0.024 ± 0.031	0.001 ± 0.011	-0.001 ± 0.029	0.006 ± 0.029
$\delta(m_{\mathrm{KT}})$	-0.049 ± 0.046	0.043 ± 0.053	0.004 ± 0.043	-0.007 ± 0.042
$\delta(\omega_{\mathrm{N1}})$	-0.003 ± 0.023	0.014 ± 0.016	-0.015 ± 0.019	0.017 ± 0.021
$\delta(\omega_{\mathrm{N2}})$	-0.017 ± 0.022	0.027 ± 0.016	0.022 ± 0.020	-0.013 ± 0.015
$\delta(\omega_{\mathrm{N3}})$	-0.004 ± 0.012	0.011 ± 0.016	0.023 ± 0.011	-0.020 ± 0.012
	Others Parameters Suppressed			

Table 9.5: Fitting for τ_{B^0}, mistag rate and Δm_d broken down by reconstructed $B^0,\overline{B}^0$ and tag side $B^0,\overline{B}^0$. The fits use the full data sample and a GExp+G resolution function in the MasterModel. The values listed are the change in central value with respect to the nominal fit.

9.6   Subsamples of Signal MC

The signal MC sample was split into four data size subsamples and each sample was fitted. A larger number of samples would be preferred to generate a reasonable histogram of results. The amount of Monte Carlo data is rather limited so the only cross check that can be done is to see if the four data size samples are consistent. The results are shown in Table 9.6.

Difference from default fit — Signal Δt fits (GExp+G) to signal MC				
	Sample1	Sample2	Sample3	Sample4
$\delta(\Delta m_d)$ (ps^{-1})	0.008 ± 0.008	-0.006 ± 0.008	-0.011 ± 0.007	-0.002 ± 0.007
$\delta(\tau_{B^0})$ (ps)	-0.011 ± 0.015	-0.008 ± 0.014	0.035 ± 0.016	-0.015 ± 0.015
$\delta(\omega_{\mathrm{LT}})$	-0.010 ± 0.006	-0.002 ± 0.006	0.009 ± 0.006	0.001 ± 0.006
$\delta(\omega_{\mathrm{KT}}^{\mathrm{offset}})$	0.014 ± 0.015	-0.004 ± 0.015	0.001 ± 0.015	-0.009 ± 0.015
$\delta(m_{\mathrm{KT}})$	-0.013 ± 0.022	0.0015 ± 0.021	-0.004 ± 0.021	0.012 ± 0.021
$\delta(\omega_{\mathrm{N1}})$	-0.001 ± 0.011	0.0085 ± 0.011	0.020 ± 0.011	-0.018 ± 0.010
$\delta(\omega_{\mathrm{N2}})$	-0.008 ± 0.010	-0.0085 ± 0.009	0.001 ± 0.009	0.006 ± 0.009
$\delta(\omega_{\mathrm{N3}})$	-0.011 ± 0.007	-0.0017 ± 0.007	0.020 ± 0.007	-0.006 ± 0.007
$\delta(\Delta\omega_{\mathrm{LT}})$	-0.008 ± 0.010	0.014 ± 0.010	-0.001 ± 0.010	-0.005 ± 0.010
$\delta(\Delta\omega_{\mathrm{KT}})$	-0.004 ± 0.008	-0.006 ± 0.008	0.006 ± 0.008	0.004 ± 0.008
$\delta(\Delta\omega_{\mathrm{N1}})$	0.005 ± 0.017	-0.015 ± 0.017	-0.011 ± 0.017	0.018 ± 0.016
$\delta(\Delta\omega_{\mathrm{N2}})$	-0.004 ± 0.015	0.000 ± 0.014	0.007 ± 0.015	-0.003 ± 0.014
$\delta(\Delta\omega_{\mathrm{N3}})$	0.004 ± 0.010	-0.014 ± 0.011	0.006 ± 0.011	0.003 ± 0.011
	Res parameters also very consistent			

Table 9.6: The signal MC sample split into four data size samples. After splitting the sample, only correctly reconstructed events are used in the four fits. The sub-samples are consistent with the full fit.

9.7   Sensitivity to NT3 tagging information

The NT3 tagging category allows the simultaneous measurement of Δm_d and τ_{B^0} to use all available events. Events with little or no tagging information are useful for the lifetime measurement but are normally thrown away to accommodate the mixing measurement.

The Q value associated with the NT3 category is extremely small, approximately 0.3%. Thus one would not expect the NT3 category to have a significant impact on the Δm_d measurement. To confirm this assumption, a fit is performed to data with the mistag rate for NT3 fixed at 50% (zero tagging power).

The change in parameters are all very minor. When the NT3 tagging information is used, the lifetime increases by 0.0022 ps and the statistical error increases by 0.002 ps larger (in quadrature); the change in Δm_d is $+0.0008$ ps^{-1} with an error that is larger by 0.0036 ps^{-1} (in quadrature). Including NT3 events is warranted just based on expected improvements to the lifetime measurement. It will also improve the Δm_d measurement, probably by a very small amount.

9.8   Subsamples by D^0 Decay Mode

The full fit to Data was performed separately for each D^0 decay mode. Table 9.7 shows the final fit values for each D^0 subsample, there is no significant variation between subsamples.

	Full Δt fits (GExp+G) to Data		
Parameter	$K^-\pi^+$	$K^-\pi^+\pi^+\pi^-$	$K\pi^+\pi^0/K_S\pi^+\pi^-$
N_{events}	14523	33621	19760
$\delta(\Delta m_d)$	0.034 ± 0.024	-0.037 ± 0.026	0.019 ± 0.025
$\delta(\tau_{B^0})$	-0.035 ± 0.027	0.017 ± 0.039	0.016 ± 0.032
$\delta(f^{\text{out}})$	0.001 ± 0.002	0.000 ± 0.003	-0.002 ± 0.002
$\delta(f_{B^+})$	0.043 ± 0.035	-0.022 ± 0.048	-0.025 ± 0.036

Table 9.7: Difference in fit results between MasterModel fits to the D^0 subsamples and the baseline fit result. Errors listed are the difference in quadrature.

9.9   Varying Δt Selection Criterion

This section looks at the choice of Δt range used in the fit. Almost all events outside ± 15 ps are outliers. A wider range is used because the resolution model is robust against such events, in fact these events give a strong handle for measuring the outlier fraction. The fit range used for data and Monte Carlo fits is ± 18 ps.

Table 9.8 shows the effect of using a smaller $|\Delta t|$ range for signal fits to signal Monte Carlo, while Table 9.9 shows the effect on the fits to data.

Signal Δt (GExp+G) fits to signal MC				
$|\Delta t|$ cut (ps)	18 (ref)	$\delta(16)(\times 10^3)$	$\delta(14)(\times 10^3)$	$\delta(10)(\times 10^3)$
Δm_d (ps^{-1})	0.4715 ± 0.0043	0.13 ± 1.3	0.1 ± 0.1	2.2 ± 1.0
τ_{B^0} (ps)	1.5523 ± 0.0091	-0.4 ± 2.8	-2.3 ± 1.2	-26.4 ± 1.9

Table 9.8: The effect of changing the Δt cut on τ_{B^0}, and Δm_d. The fits use the GExp+G resolution model and correctly reconstructed events in our large signal MC sample compared to the reference selection of $|\Delta t| < 18 ps$. The resolution and mistag parameters are consistent.

Full fits to Data		
$|\Delta t|$ cut (ps)	10	14
$\delta(\Delta m_d)$ (ps^{-1})	-0.0102 ± 0.0188	0.0016 ± 0.0020
$\delta(\tau_{B^0})$ (ps)	0.0381 ± 0.0467	-0.0023 ± 0.0047

Table 9.9: The effect of changing the Δt cut on τ_{B^0}, and Δm_d. The fits use the GExp+G resolution model on the full data sample, and since the physics parameters are blinded, the table shows the shift with respect to nominal fit result.

The number of events gained or lost is very small around the cut at 18 ps. There are a total of 16 events lost when moving from $|\Delta t| < 18$ ps to $|\Delta t| < 14$ ps in the data sample, and 114 events for the very tight cut on $|\Delta t| < 10$ ps.

The mistag rates and resolution parameters are insensitive to the Δt selection criterion. The values of τ_{B^0} and Δm_d change very little compared to the statistical errors. The changes are consistent with 0.

9.10 Varying $\sigma_{\Delta t}$ Criterion

Signal Δt fits (GExp+G) to signal MC					
$\sigma_{\Delta t}$ cut (ps)	1.4	1.6	1.8	2.0	2.2
$\delta(\Delta m_d)$ (ps^{-1})$\times 10^3$	0.1 ± 0.9	0.0 ± 0.9	**0**	0.5 ± 0.3	0.1 ± 0.8
$\delta(\tau_{B^0})$ (ps)$\times 10^3$	-0.1 ± 02.3	0.4 ± 01.8	**0**	-0.4 ± 02.2	1.3 ± 02.9

Table 9.10: The effect of changing the $\sigma_{\Delta t}$ cut on τ_{B^0} and Δm_d. The fits use the GExp+G resolution model and correctly reconstructed events in the signal MC sample. Resolution and mistag parameters are consistent across each fit.

The value of the $\sigma_{\Delta t}$ cut was determined in Sec. 5.6 based on the fact that reso-lution bias parameters scaled linearly with $\sigma_{\Delta t}$ up to a value of 1.8 ps for both resolution models (G+G+G and GExp+G). The events with a $\sigma_{\Delta t}$ larger than 1.8 ps represent less than $\approx 1.5\%$ of the total events in the sample, plus they have the largest uncertainty in the measurement of Δt, so removing these events has a minimal impact on the statistical uncertainty of the measurements. The maximum value of $\sigma_{\Delta t}$ is 2.4 ps based on a cut made in DstarlnuUser ntuples. Table 9.10 shows how parameters change with different $\sigma_{\Delta t}$ cuts. Once again the values of τ_{B^0} and Δm_d are consistent across the different sam-ples. For the Data sample, Table 9.11 shows the change in final fit results for the tighter $\sigma_{\Delta t}$ cut of 1.4 ps.

The full fit to Data with $\sigma_{\Delta t} < 1.4$ ps was done with the outlier fraction fixed to the value from a fit to Data with $\sigma_{\Delta t} < 1.8$ ps. In that case, the shift in the fitted value of τ_{B^0} is about twice that shown in Table 9.11, but is still not statistically significant.

9.11 Fixing τ_{B^0} or Δm_d to PDG2000/2002 values

Table 9.12 compares the values of τ_{B^0}, Δm_d when they are both allowed to float and when one or the other is fixed to its corresponding PDG2000 value (also the value

Full Δt fits (GExp+G) to Data	
$\sigma_{\Delta t}$ cut (ps)	1.4
$\delta(\Delta m_d)$ (ps^{-1})	0.0012 ± 0.0069
$\delta(\tau_{B^0})$ (ps)	-0.0046 ± 0.0082
$\delta(f_{B+})$	-0.004 ± 0.007
$\delta(f^{\text{out}})$	-0.0013 ± 0.0003

Table 9.11: The effect on τ_{B^0} and Δm_d of changing the $\sigma_{\Delta t}$ cut from the nominal 1.8 ps to 1.4 ps. The fits use the GExp+G resolution model for Data. Mistag parameters are consistent across each fit, but suppressed for clarity. The significant change in outlier population is plausible since the larger $|\Delta t|$ outliers are likely removed by the tighter $\sigma_{\Delta t}$ cut, and these events are the most useful for determining the outlier fraction. The central values reported are the change with respect to the baseline value, and the errors are the quadrature difference of statistical errors from the fits.

used in the Monte Carlo decay file). The fit is done using only correctly reconstructed events in the signal MC sample. The effect of fixing τ_{B^0} or Δm_d is very small. For the equivalent test on Data, the PDG2002 values are marked on the correlation plot (Fig. 11.3).

9.12 Fitting with Different Resolution Models for signal MC

The fitted parameters for the three different resolution models are shown in Table 9.13. They all do a good job of determining Δm_d, the G+G model has more trouble with determining τ_{B^0}, but this is expected since the G+G model can't model the asymmetrical bias and outlier events simultaneously. The G+G+G and GExp+G models have more freedom to capture these features. The G+G+G and GExp+G are the only serious choices for resolution models for this measurement.

9.13 Mistag rate – $\sigma_{\Delta t}$ correlation

This section looks at how sensitive the values of Δm_d and τ_{B^0} are to using the wrong value for the kaon $\sigma_{\Delta t}$ slope. The reason for considering a non-zero slope was

shown in the tagging chapter. The kaon mistag rate is the only mistag rate which shows a scaling behavior with $\sigma_{\Delta t}$. The value of the slope is varied from -2σ to $+2\sigma$. The results are shown in Table 9.14.

The value of $\omega_{\mathrm{KT}}^{\mathrm{offset}}$ changes as one would expect since the average Kaon mistag rate should stay roughly constant (Kaon mistag $= \omega_{\mathrm{KT}}^{\mathrm{offset}} + m_{\mathrm{KT}} \cdot \sigma_{\Delta t}$). All other parameters are stable and consistent with the standard fit parameters.

9.14 Varying δm Upper Limit

The sideband used for determining the combinatoric sideband extends out to $\delta m < 0.165$ GeV. This range is varied to ensure that the results do not depend on the choice of range. The range is varied from 0.155 to 0.165 GeV. The changes on the values of Δm_d and τ_{B^0} are very small and can very safely be neglected.

9.15 Summary

Figures 9.1–9.2 summarize the results of this set of cross-checks on the Monte Carlo signal sample. Figures 9.3–9.4 summarize the results of this set of cross-checks on the Data sample.

Signal Δt (GExp+G) fits to signal MC			
	$\delta(fix\ \tau_{B^0})$	$\delta(fix\ \Delta m_d)$	fit τ_{B^0}, Δm_d
Δm_d (ps^{-1})	0.0007 ± 0.00034	0.472	0.4715 ± 0
τ_{B^0} (ps)	1.548	-0.0004 ± 0.0006	1.5523 ± 0.0091
Mistag Rates			
ω_{LT}	$-0.00016 \pm 1.0e-4$	$-9.4e-05 \pm 3.7e-5$	0.0561 ± 0.0035
$\omega_{\mathrm{KT}}^{\mathrm{offset}}$	$-0.00021 \pm 1.9e-4$	$-4.2e-05 \pm 2.1e-4$	0.0672 ± 0.0084
m_{KT}	$8.0e-05 \pm 2.7e-4$	$-8.0e-05 \pm 3.1e-4$	0.1275 ± 0.0121
ω_{N1}	$-0.00015 \pm 1.8e-4$	$-9.0e-05 \pm 1.3e-4$	0.1840 ± 0.0061
ω_{N2}	$-6.0e-05 \pm 1.6e-4$	$-5.0e-05 \pm 1.4e-4$	0.3391 ± 0.0054
ω_{N3}	$0 \pm 1.2e-4$	$-1.0e-05 \pm 1.1e-4$	0.4617 ± 0.0040
Δ Mistag Rates			
$\Delta\omega_{\mathrm{LT}}$	$-9.9e-06 \pm 6.8e-5$	$3.5e-06 \pm 9.0e-5$	0.0044 ± 0.0058
$\Delta\omega_{\mathrm{KT}}$	$-4.0e-06 \pm 3.0e-5$	$-4.2e-05 \pm 3.0e-5$	-0.0140 ± 0.0045
$\Delta\omega_{\mathrm{N1}}$	$5.0e-06 \pm 4.4e-5$	$-6.0e-06 \pm 6.3e-5$	0.0120 ± 0.0098
$\Delta\omega_{\mathrm{N2}}$	$-4.6e-06 \pm 1.0e-7$	$1.6e-06 \pm 1.0e-7$	-0.0415 ± 0.0083
$\Delta\omega_{\mathrm{N3}}$	$2.0e-07 \pm 1.0e-7$	$1.0e-07 \pm 1.0e-7$	-0.0742 ± 0.0061
Resolution Parameters			
κ	0.0053 ± 0.0023	0.0008 ± 0.0017	1.1598 ± 0.0680
s^1	0.0058 ± 0.0025	0.0011 ± 0.0030	1.0017 ± 0.0216
f_{LT}^1	0.00037 ± 0.0024	-0.00031 ± 0.0010	0.8412 ± 0.0230
f_{KT}^1	0.0012 ± 0.0029	$9.0e-05 \pm 0.0011$	0.6864 ± 0.0219
f_{N1}^1	$7e-05 \pm 0.0030$	-0.00026 ± 0.0012	0.8532 ± 0.0289
f_{N2}^1	0.00085 ± 0.0029	0.00018 ± 0.0010	0.7397 ± 0.0248
f_{N3}^1	0.00093 ± 0.0026	0.00026 ± 0.0009	0.7188 ± 0.0202
f^{out}	0.00031 ± 0.0001	$2.1e-05 \pm 7.7e-5$	0.0044 ± 0.0015
b^{out}	-	-	-1.68
s^{out}	-	-	5.44

Table 9.12: The effect of fixing τ_{B^0} and Δm_d to the generator values compared to the baseline result. These fits use all correctly reconstructed signal MC events. All fits use the GExp+G resolution model; the right most column is the "standard" fit to which all other fits are generally compared. The outlier width is fixed to a value of 5.44 ps for all fits, and the width to -1.68 ps. The errors are statistical, and represent the difference in quadrature between the nominal result w/fixed τ_{B^0} or Δm_d and the result from the corresponding fixed generator value fit.

Signal Δt fits to signal MC			
	GExp+G	G+G+G	G+G
Δm_d (ps^{-1})	0.4715 ± 0.0043	0.0032 ± 0.0014	0.0030 ± 0.0009
τ_{B^0} (ps)	1.5523 ± 0.0091	-0.0017 ± 0.0007	0.0050 ± 0.0025
$\omega_{\rm LT}$	0.0561 ± 0.0035	-0.0014 ± 0.0004	-0.0011 ± 0.0004
$\omega_{\rm KT}^{\rm offset}$	0.0672 ± 0.0084	-0.0008 ± 0.0015	0.0015 ± 0.0016
$m_{\rm KT}$	0.1275 ± 0.0121	0 ± 0.0025	-0.0027 ± 0.0025
$\omega_{\rm N1}$	0.1840 ± 0.0061	-0.0010 ± 0.0006	-0.0007 ± 0.0006
$\omega_{\rm N2}$	0.3392 ± 0.0054	-0.0005 ± 0.0005	-0.0005 ± 0.0004
$\omega_{\rm N3}$	0.4617 ± 0.0040	-0.0001 ± 0.0003	-0.0002 ± 0.0003

Table 9.13: Comparison of results with three different Δt resolution models: GExp+G, G+G+G and G+G. The fits used all correctly reconstructed signal MC events. The outlier width scale factor for the GExp+G fit was fixed at 5.44 ps, and the outlier bias set to -1.68 ps. In the G+G+G fit, the width of the outlier was fixed to 8 ps and the width of the wide Gaussian was set to 2 ps. (The fit had trouble with just the outlier width fixed.) All parameters in the G+G model were floating. The G+G+G and G+G fit results are shown with reference to the GExp+G baseline result. The errors listed are the difference in quadrature between the baseline and the new result, in an effort to isolate the uncorrelated error.

Difference from default fit — Signal Δt fits (GExp+G) to signal MC				
$m_{\rm KT}$ =	0.100	0.112	0.137	0.150
$\delta(\Delta m_d)$ (ps^{-1})	0.0003 ± 0.0043	0.0001 ± 0.0043	0.0000 ± 0.0043	0.0000 ± 0.0043
$\delta(\tau_{B^0})$ (ps)	0.0002 ± 0.0087	0.0001 ± 0.0087	0.0000 ± 0.0087	0.0000 ± 0.0087
$\delta(\omega_{\rm KT}^{\rm offset})$	0.017 ± 0.003	0.009 ± 0.003	-0.007 ± 0.003	-0.015 ± 0.003
	All Resolution & Mistag Parameters Consistent			

Table 9.14: Fitting for τ_{B^0}, mistag rate and Δm_d while changing the value of the dependence of the mistag rate on Δt error for the kaon tagging category ($m_{\rm KT}$ in units of ps^{-1}). The sample is all correctly reconstructed signal MC events.
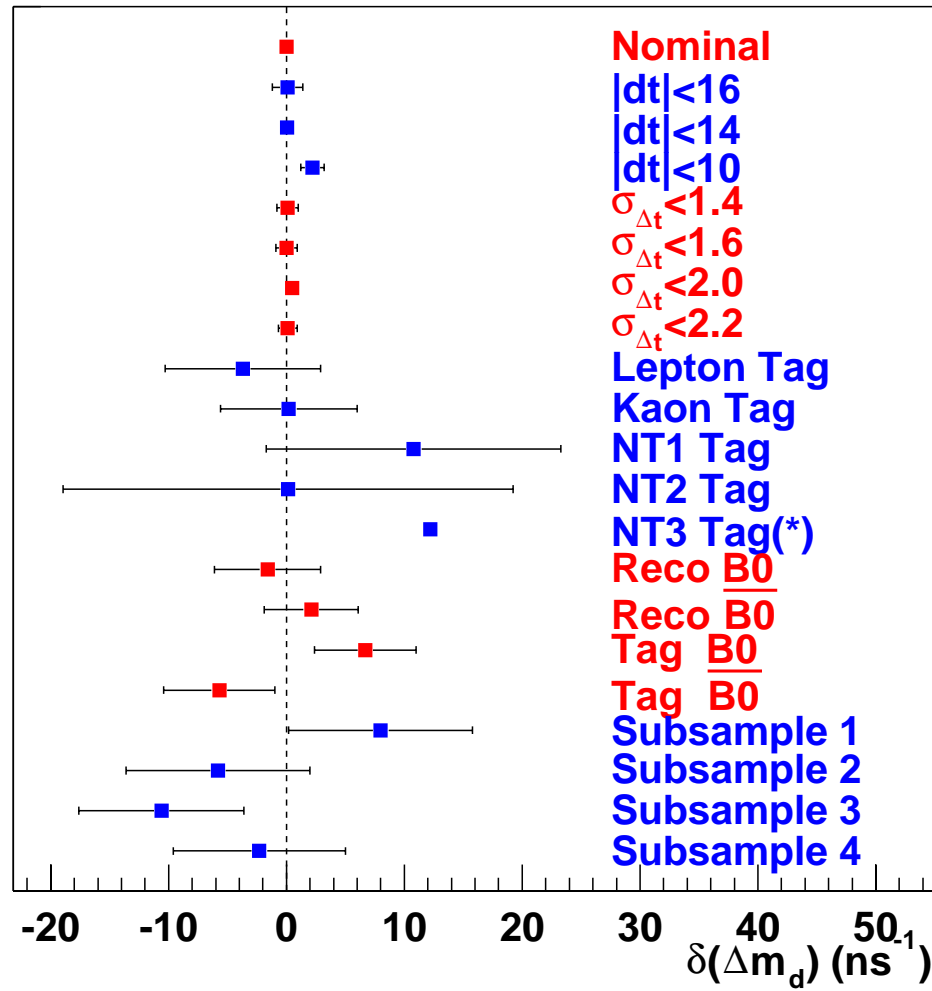
Figure 9.1: These plots show the Δm_d differences from the standard fit for various fits to the Monte Carlo signal sample. The units of the plots are ns^{-1}. The errors shown on the plots are the sample and subsample errors subtracted in quadrature as appropriate. The lines marked with (*) have had the very large errors suppressed.
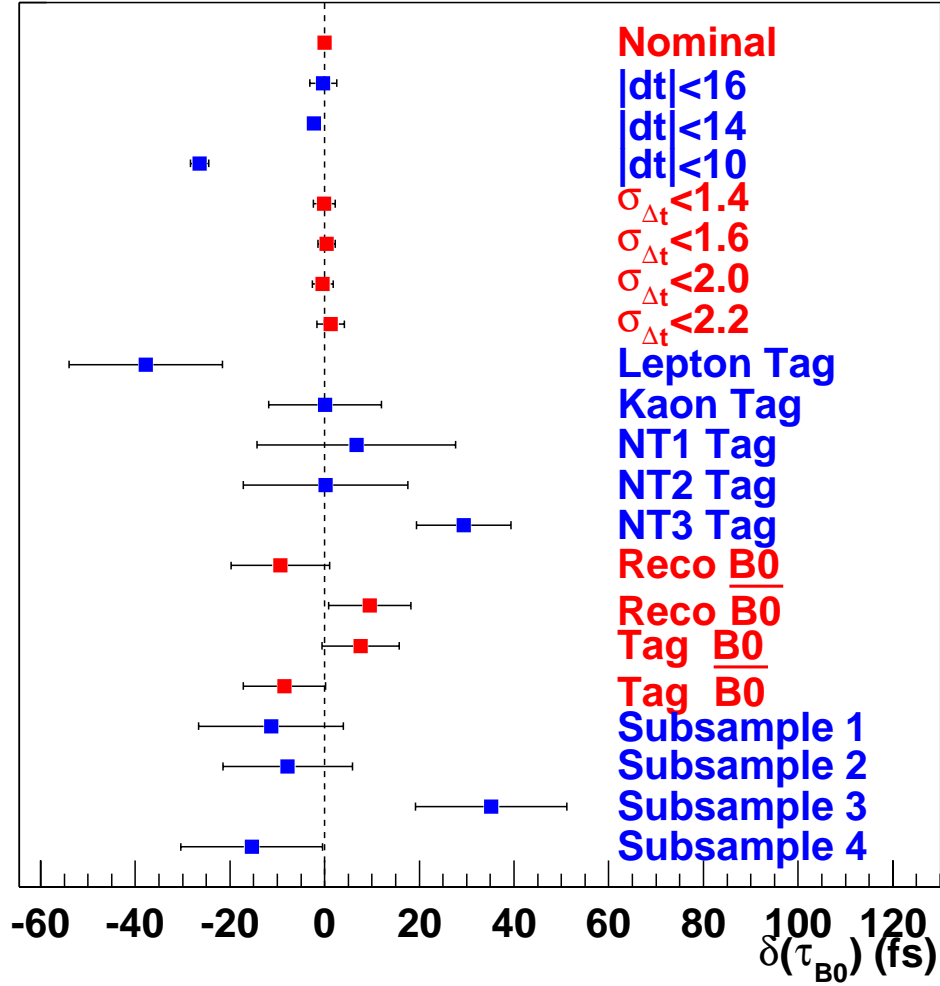
Figure 9.2: These plots show the τ_{B^0} differences from the standard fit for various fits to the Monte Carlo signal sample. The units of the plots are fs. The errors shown on the plots are the sample and subsample errors subtracted in quadrature.
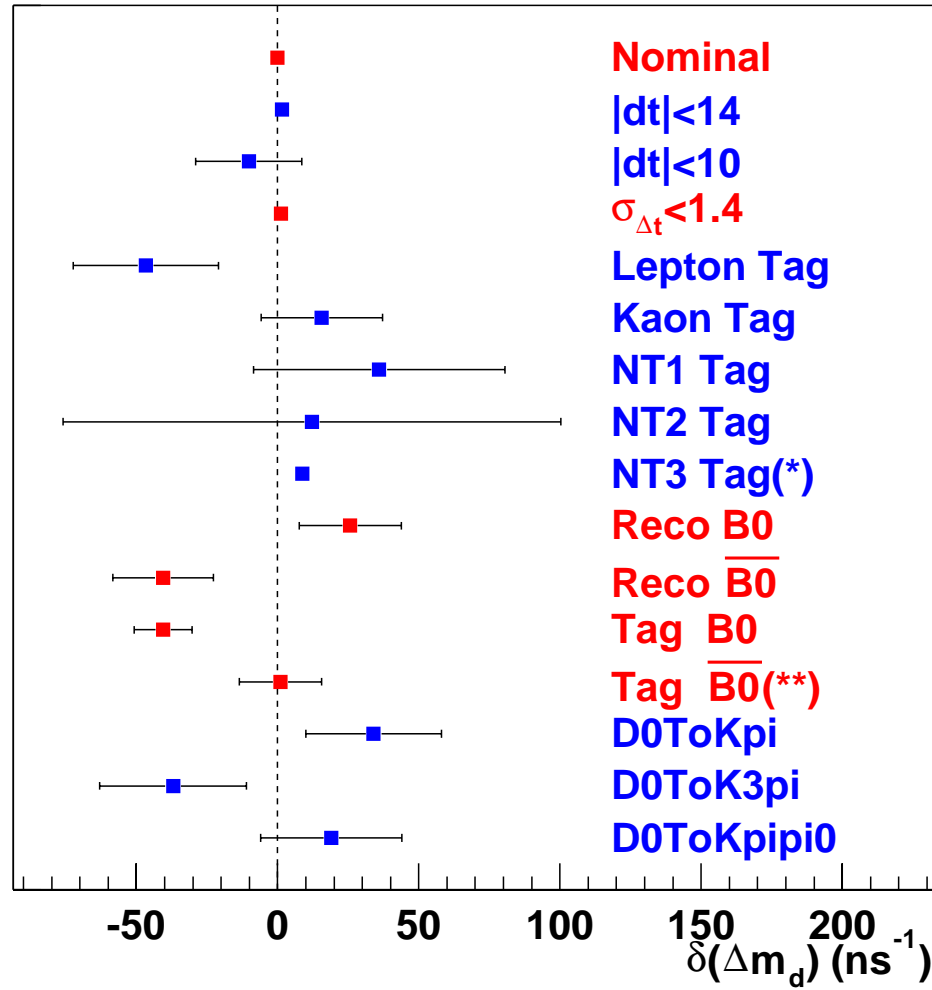
Figure 9.3: These plots show the Δm_d differences from the standard fit for various fits to the Data sample. The units of the plots are ns^{-1}. The errors shown on the plots are the sample and subsample errors subtracted in quadrature. The lines marked with an (*) indicate extremely large errors, which are suppressed. The line marked with (**) represents a fit that only converged when f_{B+} was fixed to the nominal value.
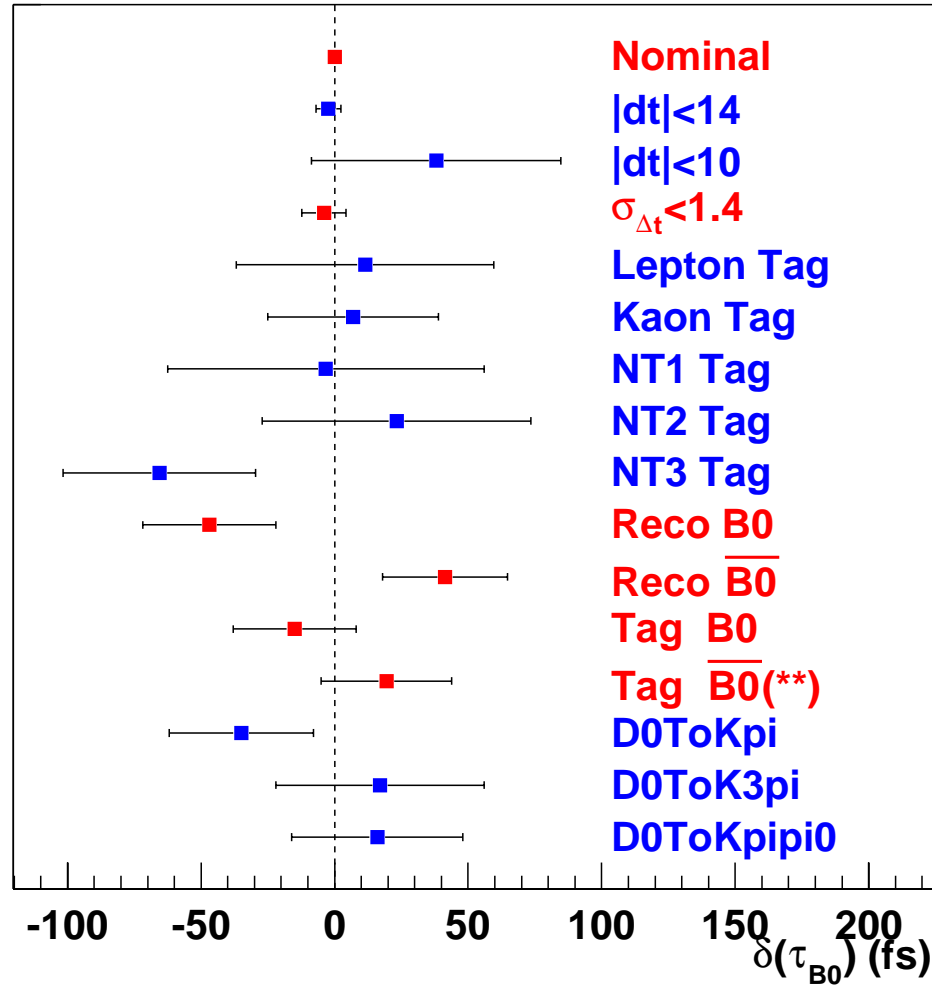
Figure 9.4: These plots show the τ_{B^0} differences from the standard fit for various fits to the Data sample. The units of the plots are fs. The errors shown on the plots are the sample and subsample errors subtracted in quadrature. The line marked with (**) represents a fit that only converged when f_{B^+} was fixed to the nominal value.

Chapter 10

Systematic Studies

This section describes the systematic error estimates for the Δm_d and τ_{B^0} measurements. The errors are estimated using both data and Monte Carlo fits.

10.1 Event Selection/Fitting Bias and Monte Carlo Correction

The event selection bias was determined by fitting correctly reconstructed events from the largest Monte Carlo sample available (the signal Monte Carlo sample) and observing any deviations from the generator values. The fit yielded $\Delta m_d = 0.4726 \pm 0.0045$ ps^{-1}, $\tau_{B^0} = 1.5529 \pm 0.0090$ ps. This does not show any significant deviation from the generated values ($\Delta m_d = 0.472$ ps^{-1}, $\tau_{B^0} = 1.548$ ps). Thus there is no statistically measurable bias in the signal Monte Carlo sample which is roughly 5 times larger than the data sample. This demonstrates that there will be no significant bias due to event selection.

The full fit using the Master Model and generic Monte Carlo does show significant bias. As discussed in Sec. 8.8, a correction of -0.022 ps for τ_{B^0} and -0.020 ps^{-1} for Δm_d was determined from the difference between a full fit to all events and a signal PDF fit to correctly reconstructed events. To be conservative, the full statistical error of the fit

is used as the systematic error, i.e., ± 0.018 ps for τ_{B^0} and $\pm 0.012\ ps^{-1}$ for Δm_d.

This uncertainty has already covered the possible event selection bias and the small bias due to ignoring the B momentum at the $\Upsilon(4s)$ frame.

10.2   Fixed Outlier Scale

A complete fit using the full data sample is repeated with 36 different fixed values of the outlier bias and outlier width, covering a wide range for each parameter (-1 to -10 ps in bias, 4 to 11 ps in width). A scatter plot of the resulting values of Δm_d and τ_{B^0} is shown in Fig. 11.2. The systematic error is assigned the value of half the full spread of lifetime and mixing results. The resulting systematic uncertainties are ± 0.005 ps for τ_{B^0} and ± 0.001 ps^{-1} for Δm_d.

10.3 z Scale

The z scale systematic error is common to many analyses which measure the time difference between B^0 decays. The difference in the z position is used to determine Δt for an event, so a bias in the z scale will directly lead to a bias in the measured Δt. A conservative estimate of the z scale uncertainty from the fully reconstructed hadronic group is less than $\pm 0.4\%$ [22]. This measurement of the time difference is very similar so the same error estimate will be used here. The error on Δm_d is less than ± 0.002 ps^{-1} while the error on τ_{B^0} is ± 0.006 ps.

10.4 PEP-II Boost Uncertainty

Since Δt is directly proportional to the measured average PEP-II boost, the errors on τ_{B^0} and Δm_d are also directly related to the uncertainty on the boost. Once again, this is a systematic error which many other measurements have already estimated. The common stated error is 0.1 % on the boost which translates into the following systematic uncertainties, $\pm 0.0005 \ ps^{-1}$ for Δm_d and ± 0.0015 ps for τ_{B^0}.

10.5 B^+ Mistag Rate, Lifetime and Fraction

The ratio of the B^+/B^0 lifetimes and mistag ratios are fixed in the final, full fit. In this section, several fits are performed to look for systematic changes in the fit results when the charged B lifetime and mistag ratios are varied. The fits use Monte Carlo and data events and vary the ratios by their known errors.

Nominal B^+/B^0 Ratios for MC	
τ_{B^+}/τ_{B^0}	1.0690 ± 0.029
$\omega_{LT}^{B^+}/\omega_{LT}^{B^0}$	1.210 ± 0.46
$\omega_{KT}^{B^+}/\omega_{KT}^{B^0}$	0.724 ± 0.12
$\omega_{N1}^{B^+}/\omega_{N1}^{B^0}$	0.740 ± 0.26
$\omega_{N2}^{B^+}/\omega_{N2}^{B^0}$	0.942 ± 0.17
$\omega_{N3}^{B^+}/\omega_{N3}^{B^0}$	1.121 ± 0.12

Table 10.1: Values for ratios of B^+ to B^0 parameters for use in the signal Δt model which describes candidates from both charged and neutral parents. The mistag rates were determined directly from Monte Carlo truth by counting.

For this study, the 20/fb sample of generic B^0 and B^+ Monte Carlo is used. Only properly reconstructed signal candidates (including those from B^+ parents) are selected. The fit is performed using the signal Δt model (GExp+G resolution) with the mistag ratio fixed to three different values, and similarly for the lifetime ratio. The nominal values

Nominal B^+/B^0 Ratios for data	
τ_{B^+}/τ_{B^0}	1.083 ± 0.017
$\omega_{\text{LT}}^{B^+}/\omega_{\text{LT}}^{B^0}$	0.54 ± 0.10
$\omega_{\text{KT}}^{B^+}/\omega_{\text{KT}}^{B^0}$	0.68 ± 0.05
$\omega_{\text{N1}}^{B^+}/\omega_{\text{N1}}^{B^0}$	0.99 ± 0.12
$\omega_{\text{N2}}^{B^+}/\omega_{\text{N2}}^{B^0}$	1.05 ± 0.07
$\omega_{\text{N3}}^{B^+}/\omega_{\text{N3}}^{B^0}$	1.12 ± 0.12

Table 10.2: Values for ratios of B^+ to B^0 parameters for use in the signal Δt model which describes candidates from both charged and neutral parents.

Signal Δt fits to generic MC				
Parameter	(τ_{B^+}/τ_{B^0})+1σ	(τ_{B^+}/τ_{B^0})-1σ	$(\omega_{B^+}/\omega_{B^0}) + 1\sigma$	$(\omega_{B^+}/\omega_{B^0}) - 1\sigma$
$\delta(\Delta m_d)$ (ps^{-1})	-0.0004	0.0004	0.0007	-0.0005
$\delta(\tau_{B^0})$ (ps)	-0.0007	0.0008	0.0001	-0.0001

Table 10.3: Fitted results from generic MC for the two physics parameters for different values of the fixed charged B properties.

for these constants is shown in Tables 10.1-10.2, the parameters are varied by one σ to

measure the associated systematic error. The results are shown in Table 10.3.

Full Δt fits to data, different lifetime ratios			
Parameter	$(\tau_{B^+}/\tau_{B^0}) + 1\sigma$	$(\tau_{B^+}/\tau_{B^0}) - 1\sigma$	diff./2
$\delta(\Delta m_d)$ (ps^{-1})	-0.00002	**0.0006**	0.0003
$\delta(\tau_{B^0})$ (ps)	-0.0016	**0.0021**	0.0019
$\delta(f_{B^+})$	-0.0025	**0.0029**	
Full Δt fits to data, different mistag ratios			
	$(\omega_{B^+}/\omega_{B^0}) + 1\sigma$	$(\omega_{B^+}/\omega_{B^0}) - 1\sigma$	diff./2
$\delta(\Delta m_d)$ (ps^{-1})	-0.00007	**-0.00023**	0.00008
$\delta(\tau_{B^0})$ (ps)	-0.0002	**0.0004**	0.0003
$\delta(f_{B^+})$	-0.0020	**-0.0026**	

Table 10.4: Fitted results from data for Δm_d, τ_{B^0} and f_{B^+} for different values of the fixed B^+ to B^0 lifetime, mistag ratio. All other fit parameters were nearly identical.

In the MasterModel fit to data, the B^+ to B^0 lifetime ratio is fixed. This ratio

varied up and down by one sigma according to PDG2002 lifetime ratio [6] and repeat the

full fit. The mixing and lifetime results are shown in Fig. 10.4. The assigned systematic

errors are half of the observed difference.

The B^+ to B^0 mistag ratios for all five tagging categories are also fixed. All five mistag ratios are moved up and down by one sigma according to Table 11.1 and repeat the full fit. The results are shown in Table 10.4. Again the assigned error is half of the observed difference.

10.6 SVT Alignment

The Δt measurement relies heavily on the precision of the B^0's vertex determination which depends on the resolution of the SVT and DCH. The vertex position is most sensitive to internal misalignments of the SVT, misalignment of the strips and wafers which record hits in the SVT. To estimate this effect, several fits are performed with different plausible alignment problems for signal Monte Carlo. The consistency of two large subsamples of real data events using different SVT alignment sets is also checked.

10.6.1 Estimate of bias using signal MC

To estimate the effect of the systematic uncertainties in the estimated positions of the SVT wafers, a 20 fb^{-1} signal Monte Carlo sample is created with different types of "misalignment". In the full signal fit, Monte Carlo truth information is used to determine the mixing status of each candidate. This improves the statistics used in the study, and focuses on the changes in Δt due to the degraded alignment scenarios.

By comparing the best-fit values for τ_{B^0} and Δm_d between the nominal "perfect" alignment and the "misaligned" sets, an estimate of the uncertainty in Δm_d and τ_{B^0} can be made. The results are shown in table 10.5.

The central value of Δm_d decreases in both misalignment scenarios, and that

	Effect of misalignment on signal Δt fits to MC		
Parameter	Nominal Value	set 1	set 2
Δm_d	0.477 ± 0.00516	-0.00316	-0.00292
τ_{B^0}	1.53 ± 0.0155	0.0057	0.0055
κ	0.812 ± 0.131	0.0452	0.00591
s^1	0.969 ± 0.0423	0.0569	0.108
f^1_{KT}	0.507 ± 0.0767	-0.0213	0.0159
f^1_{LT}	0.687 ± 0.0812	-0.0254	0.00904
f^1_{N1}	0.831 ± 0.0839	-0.0997	0.0589
f^1_{N2}	0.555 ± 0.084	0.0059	-0.0105
f^1_{N3}	0.566 ± 0.069	-0.05	0.00131
b^{out}	-0.440 ± 0.551	0.112	0.368
f^{out}	0.0291 ± 0.00498	0.000435	0.00319

Table 10.5: Parameter values and their change from nominal values for a full signal Δt fit (with Monte Carlo truth tagging information) to the signal Monte Carlo cocktail in different SVT alignment scenarios.

τ_{B^0} increases. In the case of Δm_d, the change is nearly comparable to the statistical precision of the nominal fit, indicating that misalignment is probably an important systematic uncertainty.

Based on these results, the following systematic uncertainty is assigned to the measurement of Δm_d and τ_{B^0}:

- $\delta(\tau_{B^0})$ = **-0.0056** ps

- $\delta(\Delta m_d) = -0.0030 \text{ ps}^{-1}$

10.6.2 Consistency in data between different alignments

For this study, the data sample is separated into three subsamples according to the SVT alignment scenario used in event processing and reconstruction: set "D" (29%), set "E" (53%), and "other" (18%). Each of the subsamples is fit with the full Master Model fit, independently, including the background δm analysis for evaluation of background composition/fractions. The results are shown in Table 10.6. The fitted central values for

	Full Δt fits to data	
Parameter	Set "D"	Set "E"
$\delta(\Delta m_d)$	-0.0202 ± 0.0449	-0.0001 ± 0.0308
$\delta(\tau_{B^0})$	-0.0342 ± 0.0434	-0.0083 ± 0.0227
$\delta(f^{\mathrm{out}})$	0.0034 ± 0.0035	-0.000088 ± 0.0024
$\delta(\kappa)$	-0.254 ± 0.442	-0.133 ± 0.304

Table 10.6: Results of the full fit to subsamples of the data, as grouped by SVT alignment set (here, sets "D" and "E"). The difference in central values is shown as compared to the baseline fit, but quote the statistical error on the subsample fit result.

the key parameters are statistically consistent between these two subsets of the data.

10.7   Beam Spot Position

The beam spot is used as a constraint in both the tag-side and reconstructed-side vertices. The dependence of the fitted values of τ and Δm_d on the assumed beam spot position is investigated by using the entire signal Monte Carlo and varying the beam spot position.

For this systematic study, two scenarios are investigated: one where there is a systematic shift of the beam spot position, and a second where the position is randomly smeared according to a Gaussian of fixed width. Tables 10.7- 10.8 describe the effect observed in the final parameter values after performing the full fit (including resolution functions) on the signal sample. For these results, the GExp resolution model was used exclusively.

The lifetime, τ_{B^0}, is very robust with respect to movements of the beam spot position. The resolution function model adapts easily even to the most significant smearing and offset of the beam spot position. The mixing frequency, Δm_d, appears more sensitive, and the measured dilution parameters change slightly as the beam spot is changed.

Change in fit results to signal MC w/ beam spot variations					
		Beam spot Variation [μm]			
Parameter	nominal err	shift 10	shift 30	shift 80	shift/smear 20
Δm_d	± 0.00435	-0.00054	-0.00129	-0.00202	0.00008
τ_{B^0}	± 0.00878	-0.0042	-0.0001	0.0012	-0.0024
mistag rates	All values very consistent				
κ	± 0.0847	-0.004	-0.0125	-0.0853	0.0307
f^1_{KT}	± 0.0304	-0.00371	-0.0106	-0.0214	0.0149
f^1_{LT}	± 0.0278	0.00203	0.00343	0.00606	0.0167
f^1_{N1}	± 0.034	0.00443	0.00566	0.0005	0.0224
f^1_{N2}	± 0.0312	0.00372	-0.00967	-0.0162	0.0169
f^1_{N3}	± 0.0271	-0.00274	-0.00909	-0.0177	0.00023
b^{out}	± 0.699	-0.124	-0.476	-0.821	-0.269
f^{out}	± 0.00318	-0.000225	-0.00101	0.000076	-0.000683
s^1	± 0.0227	0.0159	0.019	0.0997	0.0203

Table 10.7: Difference in final parameter values from the full signal fit to signal Monte Carlo samples for different variations applied to the beam spot position. The beam spot position is systematically shifted before applying reco- and tag-side vertexing.

To investigate this further, another series of signal Δt fits is performed to the different Monte Carlo samples where we use Monte Carlo truth information to determine flavor tagging status. This increases the statistical precision by nearly a factor of two, results are shown in table 10.9.

By studying the z-residuals of each vertex (reco and tag) separately, it can be seen that this uncertainty in the beam spot constraint affects each vertex about equally. From figures similar to Fig. 5.4 for each of the transformed-beam spot datasets, table 10.10 is produced, which records the effect of the beam spot transformation on the z-residual distribution's width and bias.

Based on these observations, and knowing that (a) the difference in estimated beam spot position between different algorithm is than 10 μm and (b) that average resolution of the beam spot is about 30 μm, the systematic error due to uncertainties in the beam spot position is estimated to be

Change in fit results to signal MC w/ beam spot variations				
	Beam spot Variation [μm]			
Parameter	smear 20	smear 40	smear 80	shift/smear 80
Δm_d	0.00053	-0.00022	-0.00295	-0.00577
τ_{B^0}	-0.005	-0.0036	-0.0029	0.0024
mistag rates	All values consistent.			
f^1_{KT}	0.00685	0.0104	-0.0126	-0.0271
f^1_{LT}	0.00078	0.00036	-0.0197	-0.0108
f^1_{N1}	0.0158	0.00839	0.0124	-0.00667
f^1_{N2}	-0.00742	0.0107	-0.0265	-0.0328
f^1_{N3}	0.00065	0.00945	-0.0107	-0.0189
b^{out}	0.383	0.303	0.0398	-0.461
f^{out}	-0.000712	-0.000468	0.000463	0.00187
s^1	0.0248	0.0314	0.0947	0.146

Table 10.8: Difference in final parameter values from the full signal fit to the signal Monte Carlo samples for different variations applied to the beam spot position. Here the beam spot position is randomly smeared by sampling from a Gaussian of the indicated width, before applying reco- and tag-side vertexing.

- $\delta(\tau_{B^0}) = \pm 0.005$ ps

- $\delta(\Delta m_d) = \pm 0.001$ ps^{-1}

Parameter	Nominal Value	Fits to signal MC using truth-tagging				
		shift 10	shift 80	smear 20	smear 80	both 80
Δm_d	0.472 ± 0.00216	0.00022	-0.00106	0.00008	-0.00026	-0.00262
τ_{B^0}	1.54 ± 0.00686	-0.0017	0.0045	-0.0022	-0.0012	0.0053
κ	0.928 ± 0.0668	-0.0266	-0.0482	-0.00771	-0.00666	-0.0341
s^1	1.02 ± 0.0182	0.0117	0.085	0.0177	0.0874	0.14
b^{out}	-1.48 ± 0.357	-0.151	-0.405	0.192	-0.0836	-0.52

Table 10.9: Fit results for select parameters from signal Δt fits to Monte Carlo samples with a shifted and/or smeared beam spot constraint. As described in the text, Monte Carlo truth information was used to determine the flavor status of each candidates, thereby improving the statistical precision on the other physics parameters. The first column of numbers shows the fit results to the nominal signal Monte Carlo sample and their statistical error, the other columns show the change in central value.

Variation	Size	Characteristics of z-vtx residuals from signal MC			
		z_{reco} Bias	z_{reco} RMS	z_{tag} Bias	z_{tag} RMS
nominal	0	0.05 ± 0.3	77.5 ± 0.2	34.4 ± 0.5	156.1 ± 0.4
shift	10	0.07	77.4	34.6	156.8
shift	80	1.92	85.5	36.7	162.2
smear	20	0.09	77.9	34.1	156.8
smear	80	0.03	85.0	34.8	161.6

Table 10.10: Shape parameters for tag- and reco- vertex z-residual distributions of correctly reconstructed signal Monte Carlo under several different beam spot transformation scenarios. [All units are μm.] As expected, systematic offsets of the beam spot increase the bias and RMS of the z-residual for each vertex separately, while smearing only degrades the rms of the residuals.

10.8   Background Fractions

This section examines the sensitivity of Δm_d and τ_{B^0} to the calculated background fractions (combinatoric, fake lepton, uncorrelated lepton and continuum events). These parameters are fixed before the final fit is performed. The statistical uncertainty on these measured background fractions need to be accounted for. The complicated nature of the δm fits requires a novel technique to properly understand the impact of the measured background fractions on the values of Δm_d and τ_{B^0}.

The background fraction are calculated using the δm fits to peak and background shapes for the 360 possible different categories plus background fractions in each category which describes the number of signal and background events in each category. These sets of parameters are varied jointly and randomly, taking into account the individual errors on each result as well as the correlation between any two results (via the covariance matrix). This will generate an new set of background fractions. The full fit is run with the new background fractions and all other parameters fixed to baseline values except Δm_d and τ_{B^0}. This was done due to the long fit times when all parameters are allowed to float. This procedure was done 111 times to produce a reasonable set of data points. An additional 20 fits were done with all parameters floating to ensure that the larger sample gives a reasonable estimate. The results are shown in Fig. 10.1 and Tab. 10.11. When floating only the two physics parameters, the statistical error from the fit is ± 0.012 ps^{-1} on Δm_d and ± 0.017 ps on τ_{B^0}.

There were two different perturbations performed during this study which affect the background fractions (varying the δm shape fit values, and varying the yields fit results used to compute the background fractions). Each perturbation is treated separately, it was found that the dominant source of spread in the fitted values of Δm_d and

Full fits to data, floating only Δm_d, τ_{B^0}		
	$\delta(\Delta m_d)$ (ps^{-1})	$\delta(\tau_{B^0})$ (ps)
Mean	-0.0013 ± 0.0003	-0.0006 ± 0.0003
RMS	0.0029 ± 0.0002	0.0032 ± 0.0002

Table 10.11: Characteristics of the distribution of fitted results for Δm_d and τ_{B^0} (with respect to the baseline) from the 111 fits with perturbed background fraction when floating only the two physics parameters.

τ_{B^0} was due to the perturbation of the yield fits.

An additional cross-check used 20 different sets of perturbed parameters in a full fit to data. Although the precision is poor, the rms of the distributions of Δm_d and τ_{B^0} values are consistent with the 111 two-parameter fits, and the bias with respect to the baseline fit is very small. The fitted Δm_d values from the 20 different fits using perturbed background parameters yield an average shift (wrt the nominal fit) of -0.002 with an rms of 0.0012 ps^{-1}. Likewise for τ_{B^0}, the distribution of fitted values have an average of 0.0016 and an rms of 0.0033 ps. This serves as a cross-check for the systematic errors determined in this section.

Based on these results, and recognizing that $\sim$100 tests give us a fractional precision of order 7% on the rms of the distribution, we assign the following systematic errors:

- $\delta(\Delta m_d) = \pm 0.0029$ ps^{-1}

- $\delta(\tau_{B^0}) = \pm 0.0032$ ps.

10.9   Fitting with Different Resolution Model

The data is fit using the G+G+G resolution model instead of the GExp. The outlier parameters are same, ($\sigma_{\text{out}} = 6$ ps and $\mu_{\text{out}} = -5$ ps). All other parameters are

allowed to float. This fit has 74 parameters (the G+G+G model has two more parameters than the GExp model). Compared to the GExp fit, the lifetime shifts by $+0.0034$ ps and Δm_d shifts by $-0.0009 \ ps^{-1}$.

10.10 Background Δt Models

The background PDFs have all their parameters floating to minimize any kind of fitting bias, however, it is possible to introduce a bias with a poor choice of PDFs. To test the sensitivity of Δm_d and τ_{B^0} to the choice of background PDFs, the largest background (combinatoric) is altered to have only a lifetime component (no mixing term). This clearly should have more impact on Δm_d and τ_{B^0} than any other background PDF choice.

After performing the full fit with this alteration, the lifetime shifts by $+0.0063$ ps and Δm_d shifts by $-0.0012 \ ps^{-1}$. The charged B fraction, f_{B^+}, also changes from 8% to 6%. This shift is used to estimate the systematic uncertainty.

10.11 Dependence on Known PID Efficiencies

There are two places to worry about PID efficiencies. One is that some event selection bias could enter due to the use of Kaon PID for finding the D^0 and lepton PID used in finding the primary lepton. Since there is no measurable event selection bias, these concerns can be put aside. A second place that PID efficiencies enter this measurement is in the determination of the fake lepton background fraction. This turns out to have a small effect on the fake leptons, which are significantly smaller than the combinatoric background. The effect is completely negligible.

10.12 Summary

Table 10.12 summarizes the systematic errors for τ_{B^0} and Δm_d, and the sum in quadrature.

Source	$\delta(\Delta m_d)$ (ps^{-1})	$\delta(\tau_{B^0})$ (ps)	Note
Selection and fit correction	± 0.0123	± 0.0178	Stat error on generic MC fit
z scale	± 0.0020	± 0.0060	0.4% recipe
PEP-II boost	± 0.0005	± 0.0015	0.1% recipe
SVT alignment	± 0.0030	± 0.0056	uses signal MC
Beam spot position	± 0.0010	± 0.0050	Move/smear BS in sig MC
Bkg / signal prob.	± 0.0029	± 0.0032	Vary δm fits
Bkg Δt models	± 0.0012	± 0.0063	Vary comb. bkg model
Fixed B^+/B^0 lifetime ratio	∓ 0.0003	± 0.0019	Vary lifetime ratio by $\pm 1\sigma$ (PDG2002)
Fixed B^+/B^0 mistag ratio	∓ 0.0001	∓ 0.0003	Vary mistag ratios by $\pm 1\sigma$ (BAD119)
Fixed signal outlier shape	± 0.0010	± 0.0054	Vary outlier width & bias
Signal resolution model	± 0.0009	± 0.0034	G+G+G vs GExp+G
Total syst. error	± 0.013	± 0.022	

Table 10.12: Summary of systematic uncertainties on the two physics parameters, τ_{B^0} and Δm_d.
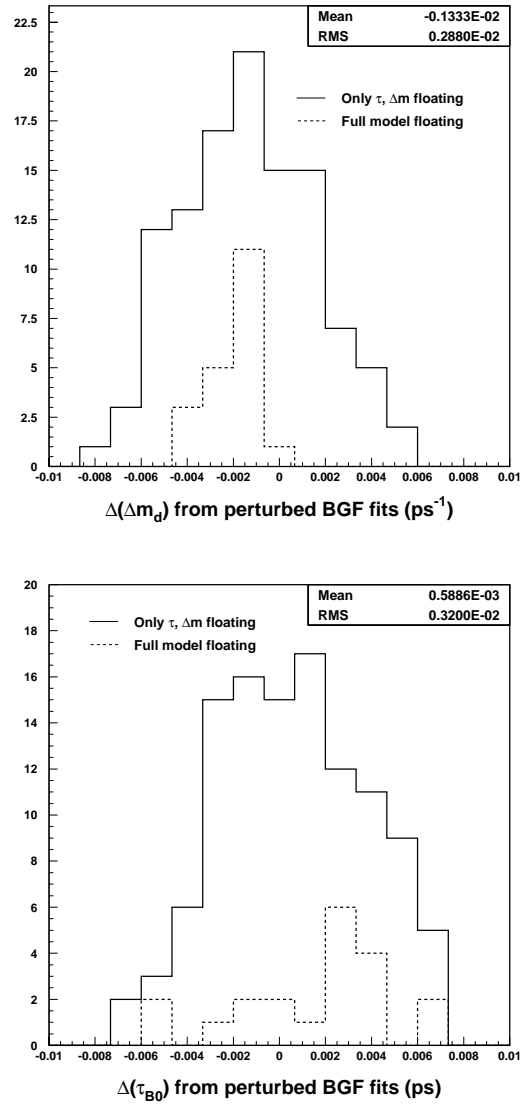
Figure 10.1: Histograms of the spread in fitted results (wrt the baseline blinded value) from data using 111 different statistical variations of the input δm parameters and floating only τ_{B^0} and Δm_d in the final Δt fits. The superimposed shaded histograms show the distribution of fit results for the 20 fits where the entire model was floating in the fit.
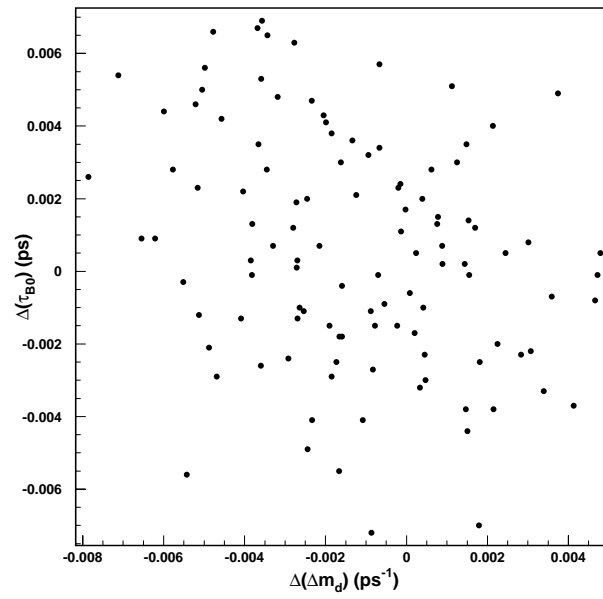
Figure 10.2: Scatter plot of pairs of $(\tau_{B^0}, \Delta m_d)$ values (wrt baseline fit) from the 111 2-parameter floating Master fits to the data, each using a different set of perturbed background fractions in the fit. We observe no induced correlation on τ_{B^0} and Δm_d due to varying the δm fit results and background fractions.

Chapter 11

Fit Results

11.1 Outlier determination

The outlier term in the resolution function accounts for events which are incorrectly reconstructed and whose per-event-error estimate is poor. Through experience, it has been determined that floating all parameters of the outlier Gaussian leads to unstable and unreliable results for the outlier parameters. This is due mainly to the very small number of outliers in the sample and sizable correlations to other parameters. Currently all other BaBar measurements that use similar resolution functions typically fix the outlier bias and scale factors. This greatly increases the stability of the fit and reduces the amount of time required for each fit. This improvement comes with a relative small cost, the fixed outlier term makes a small contribution to the systematic error. The only term associated with the outliers which floats is the outlier fraction.

The outlier width and bias are determined by performing the full fit with 36 different *fixed* values of the outlier bias and outlier width, covering a wide range for each parameter (-1 to -10 ps in bias, 4 to 11 ps in width). Figure 11.1 shows the negative-ln-

likelihood surface as a function of outlier bias and width. The negative-ln-likelihood has

been offset so that the minimum value over the plot is zero.
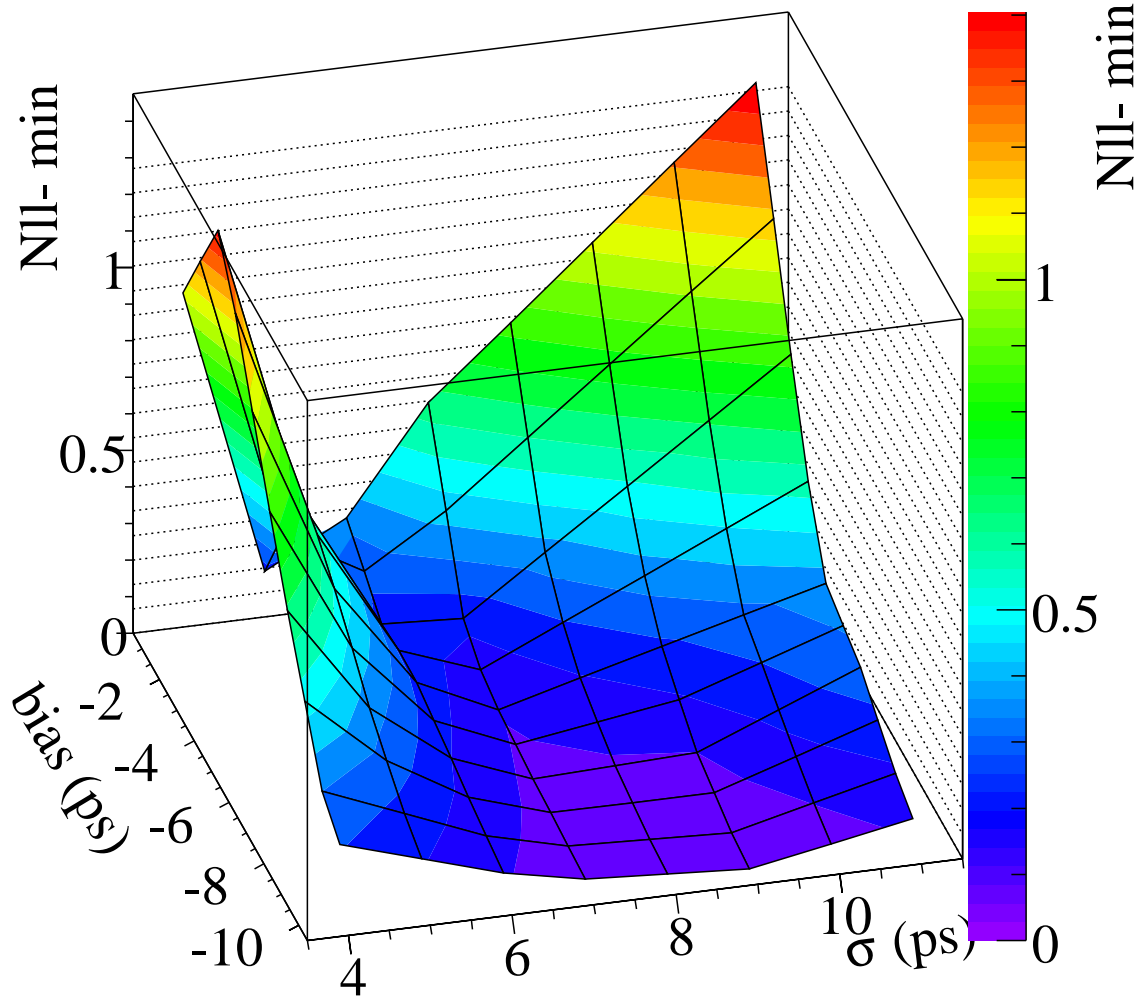
Figure 11.1: The negative-ln-likelihood surface in the space of outlier bias and outlier width for the outlier Gaussian of the signal resolution model in Data. The vertical scale has been offset so that minimum is at 0. The surface is constructed from 36 fits to the Data with different fixed values of the signal outlier parameters.

A scatter-plot of the resulting values of Δm_d and τ_{B^0} is shown in Fig. 11.2. Note

that the spread in values of τ_{B^0} and Δm_d is small compared to the statistical uncertainty

on each quantity. This is consistent with the fact that the change in ln-likelihood is small

as the outlier bias and width are varied (see Fig. 11.1).

In order to choose an outlier bias and width for the default fit to data, a point near the middle of the cluster of points in the τ_{B^0} versus Δm_d scatter-plot is selected (Fig. 11.2). The final values are a bias of -5 ps and a width of 6 ps.
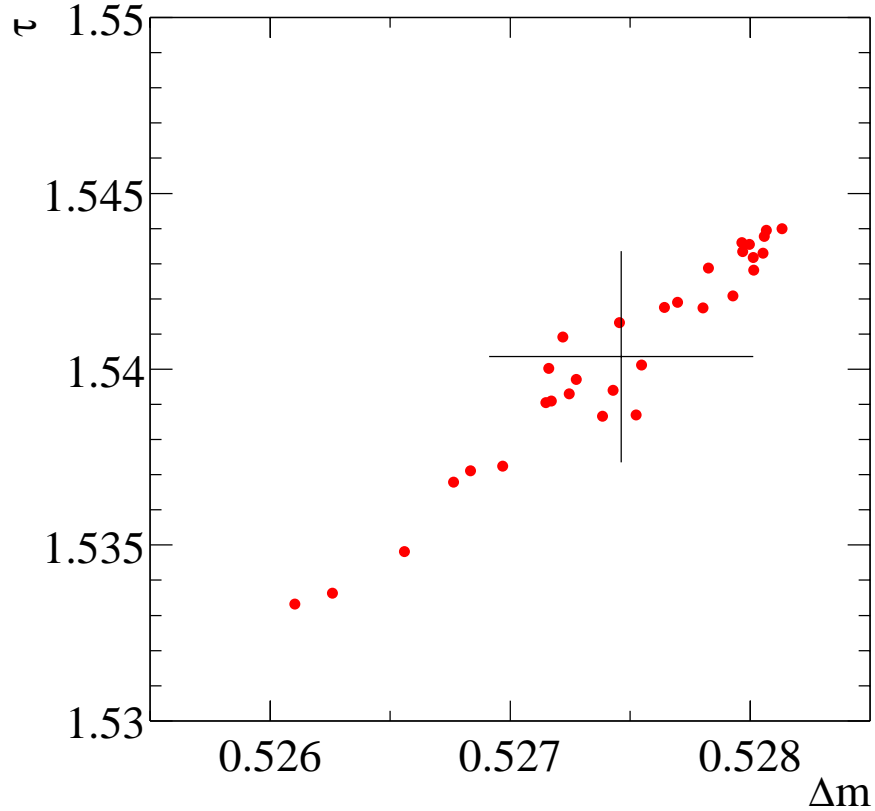
Figure 11.2: Scatter plot of the set of (blinded) Δm_d, τ_{B^0} values obtained from data for 36 fits with different fixed values of the outlier bias and width in the range -1 to -10 ps for bias and 4 to 11 ps for width. The cross indicates the mean and rms of the Δm_d, τ_{B^0} distribution. Note that the range of the τ_{B^0} (y) axis is slightly less than one statistical error bar, and that the range of the Δm_d (x) axis is less than 1/5 of a statistical error bar.

The systematic uncertainty is given by half the full spread of lifetime and mixing results. The resulting systematic uncertainties are 0.005 ps for τ_{B^0} and 0.001 ps^{-1} for Δm_d.

11.2 Full unblinded fit to data

The full master fit has a total of 72 floating parameters: 22 in signal model, 24 in combinatoric model and 26 in peaking background models. The ratio of B^+ to B^0 lifetimes is fixed at the central value from PDG2002, 1.083 ± 0.017. The ratios of B^+ to B^0 mistag fractions are fixed at the central values extracted from BAD119 [21]. They are summarized in Table 11.1.

Category	ratio
LTag	0.54 ± 0.10
KTag	0.68 ± 0.05
NT1	0.99 ± 0.12
NT2	1.05 ± 0.07
NT3	$1.12 \pm 0.12^*$

Table 11.1: B^+ to B^0 mistag fraction ratios extracted from BAD119 [21], except the NT3 category, which is not available. We calculate the ratio for NT3 from generic Monte Carlo sample.

The fitted signal Δt model parameters are shown in Table 11.2, the combinatoric Δt model parameters in Table 11.3, and the three peaking background Δt model parameters in Table 11.4. The one-sigma contour for τ_{B^0} vs. Δm_d is shown in Fig. 11.3. The correlation coefficient is $\rho(\Delta m_d, \tau_{B^0})$= -0.22. This number along with other large correlations between τ_{B^0} or Δm_d and a few other parameters are shown in Table 11.5.

Figure 11.4 and 11.5 show the Δt and the asymmetry distributions of unmixed and mixed events in signal sample and the projections of the model.

The charged B fraction in data from this final fit is $(8.2 \pm 2.9)\%$. Although the selection criteria are not exactly the same, it is interesting to compare this result with the charged B fraction found for the $B \to D^* l \nu$ sample described in the $\sin 2\beta$-mixing Phys. Rev. D journal paper: $(4.5 \pm 0.3 \pm 2.2)\%$. This result was determined from a fit

Final Fit to Data, Signal Model Parameters					
parameter	value		value		value
Δm_d (ps^{-1})	0.512 ± 0.018	f_{B+}	0.082 ± 0.029	s^1	1.201 ± 0.063
τ (ps)	1.545 ± 0.023	$\omega_{\rm LT}$	0.071 ± 0.015	κ	0.86 ± 0.17
-	-	$\omega_{\rm KT}^{\rm offset}$	0.002 ± 0.024	$f^1_{\rm LT}$	0.72 ± 0.10
-	-	$m_{\rm KT}$	0.229 ± 0.036	$f^1_{\rm KT}$	0.609 ± 0.088
-	-	$\omega_{\rm N1}$	0.212 ± 0.020	$f^1_{\rm N1}$	0.69 ± 0.13
-	-	$\omega_{\rm N2}$	0.384 ± 0.018	$f^1_{\rm N2}$	0.70 ± 0.10
-	-	$\omega_{\rm N3}$	0.456 ± 0.012	$f^1_{\rm N3}$	0.723 ± 0.078
-	-	$\Delta\omega_{\rm LT}$	-0.001 ± 0.022	$f^{\rm out}$	0.0027 ± 0.0017
-	-	$\Delta\omega_{\rm KT}$	-0.024 ± 0.015	$b^{\rm out}$ (ps)	-5.000
-	-	$\Delta\omega_{\rm N1}$	-0.098 ± 0.032	$s^{\rm out}$ (ps)	6.000
-	-	$\Delta\omega_{\rm N2}$	-0.112 ± 0.028	-	-
-	-	$\Delta\omega_{\rm N3}$	-0.023 ± 0.019	-	-

Table 11.2: Default fit result for Data — signal model parameters.

to the cos_{D*l} distribution in data, with shapes taken from Monte Carlo. Given that the errors on the two results are not correlated, the difference is $(3.7 \pm 3.7)\%$. Therefore, the two results are consistent.

Fig. 11.7 shows how the errors and correlation of τ_{B^0} and Δm_d change if other parameters are fixed at their best fit value. A series of fits is performed, fixing all parameters at the values obtained from the default fit, except (a) Δm_d and τ_{B^0}, (b) Δm_d, τ_{B^0} and all mistag fractions in signal model, (c) Δm_d, τ_{B^0} and f_{B+}, (d) Δm_d, τ_{B^0}, f_{B+} and all mistag fractions in signal model, (e) all parameters in signal Δt model. The resulting errors on Δm_d, τ_{B^0} and their correlations are shown in Table 11.6. The one-sigma contours for these fits and the default fit are shown in Fig. 11.7.

The error on τ_{B^0} changes very little until the signal resolution function parameters are floated. Floating the background parameters adds a very small contribution to the error. The contribution from the charged B fraction and mistag fractions to τ_{B^0} error is negligible. On the other hand, the charged B fraction changes the error on Δm_d

Final Fit to Data, Combinatoric Δt Model Parameters			
parameter	value	parameter	value
Δm^{comb}	0.422 ± 0.020	b^1_{LNT12}	-0.056 ± 0.021
τ^{comb}	1.234 ± 0.024	b^1_{KNT3}	-0.104 ± 0.012
$\omega^{\mathrm{osc}}_{\mathrm{Fake;LTag}}$	0.434 ± 0.026	$f^{\mathrm{osc}}_{\mathrm{Off}}$	0.099 ± 0.035
$\omega^{\mathrm{osc}}_{\mathrm{Fake;KTag}}$	0.128 ± 0.017	$f^{\mathrm{osc}}_{\mathrm{SS}}$	0.385 ± 0.023
$\omega^{\mathrm{osc}}_{\mathrm{Fake;NT1}}$	0.443 ± 0.032	$f^{\mathrm{osc}}_{\mathrm{NoLept}}$	0.526 ± 0.019
$\omega^{\mathrm{osc}}_{\mathrm{Fake;NT2}}$	0.507 ± 0.024	$f^{\mathrm{osc}}_{\mathrm{OneLept}}$	0.894 ± 0.022
$\omega^{\mathrm{osc}}_{\mathrm{Fake;NT3}}$	0.589 ± 0.020	$f^{\mathrm{osc}}_{\mathrm{TwoLept}}$	1.000
$\omega^{\mathrm{osc}}_{\mathrm{Lept;LTag}}$	0.156 ± 0.023	s^1_{OffSS}	1.337 ± 0.024
$\omega^{\mathrm{osc}}_{\mathrm{Lept;KTag}}$	0.297 ± 0.010	$s^1_{\mathrm{On;OS}}$	1.259 ± 0.021
$\omega^{\mathrm{osc}}_{\mathrm{Lept;NT1}}$	0.377 ± 0.025	f^1	0.9666 ± 0.0071
$\omega^{\mathrm{osc}}_{\mathrm{Lept;NT2}}$	0.401 ± 0.016	b^{out}	-0.98 ± 0.24
$\omega^{\mathrm{osc}}_{\mathrm{Lept;NT3}}$	0.491 ± 0.011	s^{out}	4.60 ± 0.42
ω^{pmt}	0.3960 ± 0.0093	-	-

Table 11.3: Default fit result for Data — Δt model parameters for combinatoric background, $G^{(\mathrm{comb})}$.

the most. The contributions from floating the mistag fractions, resolution functions and

background models are relatively small.

Final Fit to Data, Peaking BG Model Parameters					
cont. par.	value	fake par.	value	unco. par.	value
ω_{KT}	0.083 ± 0.028	Δm^{fake}	0.444 ± 0.031	τ^{unco}	1.07 ± 0.17
ω_{LNT}	0.457 ± 0.040	τ^{fake}	1.416 ± 0.060	$\omega^{\text{life}}_{\text{KNT}}$	0.71 ± 0.13
$b^1{}_{\text{OS}}$	0.04 ± 0.11	$\omega^{\text{osc}}_{\text{LT}}$	0.248 ± 0.038	$\omega^{\text{life}}_{\text{LT}}$	0.9998
$b^1{}_{\text{SS}}$	-0.08 ± 0.13	$\omega^{\text{osc}}_{\text{KT}}$	0.142 ± 0.029	ω^{pmt}	0.00000 ± 0.00064
s^1	1.300 ± 0.090	$\omega^{\text{osc}}_{\text{NT1}}$	0.342 ± 0.044	f^{life}	0.877 ± 0.085
f^1	0.909 ± 0.040	$\omega^{\text{osc}}_{\text{NT2}}$	0.377 ± 0.039	b^1	-0.30 ± 0.18
-	-	$\omega^{\text{osc}}_{\text{NT3}}$	0.512 ± 0.042	f^1	1.000
-	-	$\omega^{\text{pmt}}_{\text{KNT3}}$	0.46 ± 0.14	s^1	1.34 ± 0.24
-	-	$\omega^{\text{pmt}}_{\text{LNT12}}$	0.64 ± 0.15	-	-
-	-	f^{osc}	0.852 ± 0.046	-	-
-	-	$b^1{}_{\text{KNT3}}$	-0.208 ± 0.062	-	-
-	-	$b^1{}_{\text{LNT12}}$	-0.078 ± 0.065	-	-
-	-	f^1	0.944 ± 0.023	-	-
-	-	s^1	1.159 ± 0.091	-	-

Table 11.4: Default fit result for Data — Δt model parameters for peaking backgrounds. The outlier bias (μ_{out}) and scale factor (σ_{out}) for these peaking background are shared with combinatoric background parameters.

Δm_d **global correlation**	**0.74**
τ_{B^0} **global correlation**	**0.69**
$\rho(\Delta m_d, \tau_{B^0})$	-0.22
$\rho(\Delta m_d, f_{B^+})$	**0.58**
$\rho(\tau_{B^0}, \sigma^1_{\text{sig}})$	-0.49
$\rho(\tau_{B^0}, f^{\text{out}}_{\text{sig}})$	-0.26

Table 11.5: Global correlation coefficients for Δm_d and τ_{B^0} from the full fit to Data. Correlation coefficients for pairs of key parameters.

floating pars	$\sigma(\Delta m_d)$ (ps^{-1})	$\sigma(\tau_{B^0})$ (ps)	$\rho(\Delta m_d, \tau_{B^0})$
$\Delta m_d, \tau_{B^0}$	**0.0124**	**0.0166**	-0.072
$\Delta m_d, \tau_{B^0}$, mistag	**0.0132**	**0.0166**	-0.057
$\Delta m_d, \tau_{B^0}, f_{B^+}$	**0.0163**	**0.0169**	-0.161
$\Delta m_d, \tau_{B^0}, f_{B^+}$, mistag	**0.0174**	**0.0169**	-0.149
all signal Δt	**0.0178**	**0.0229**	-0.224
default fit	**0.0182**	**0.0230**	-0.216

Table 11.6: Comparison of errors on Δm_d and τ_{B^0} and their correlation coefficients in full fits to Data in which different sets of parameters are floating.
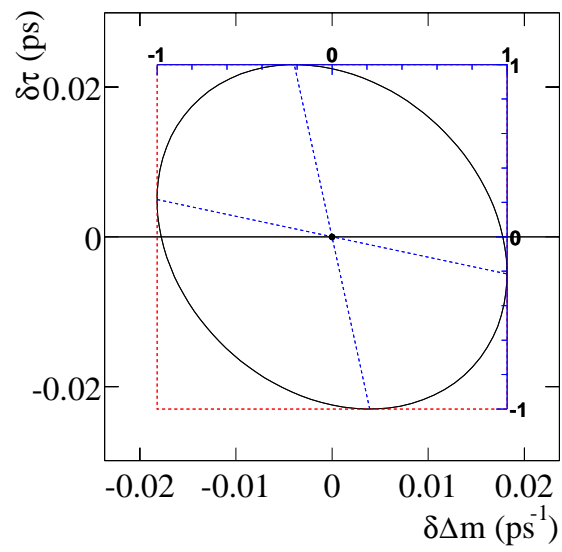
Figure 11.3: The one-sigma contour for τ_{B^0} *vs.* Δm_d from the full fit to Data. The correlation coefficient is -0.22.
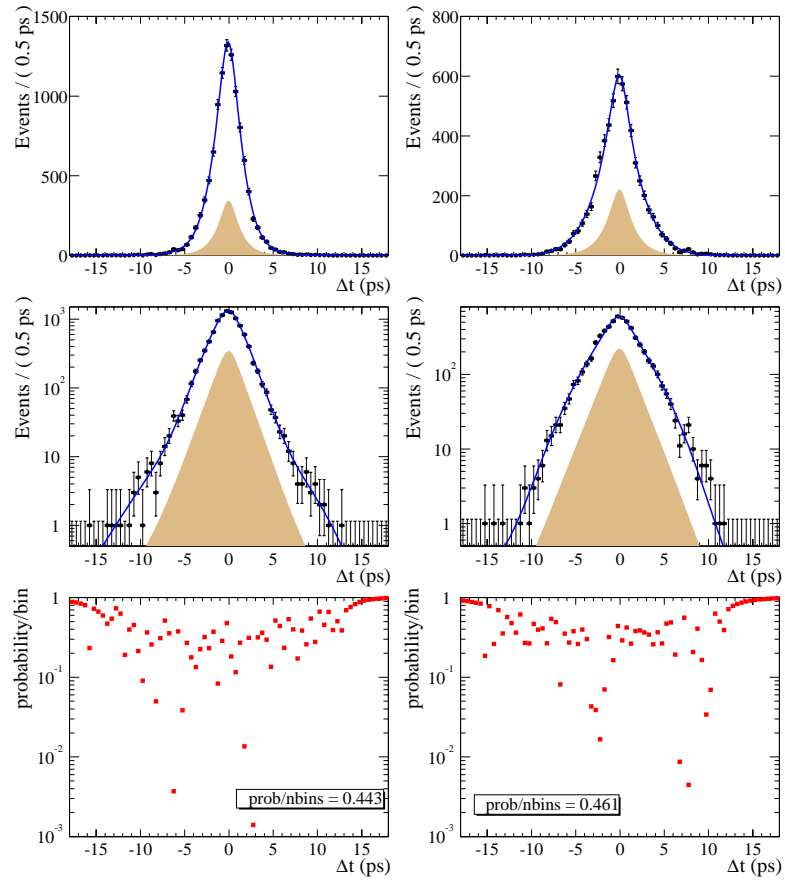
Figure 11.4: The Δt distribution of unmixed and mixed events in signal sample of Data within the peak region $143 < \delta m < 148$ MeV, and the projection of the model. The left hand plots are for unmixed events, the right for mixed events; the middle row is simply a log-y plot, The superimposed shaded areas on the upper plots show the background contribution to the distributions. The bottom row is the Poisson probability of observing n events or smaller (larger) if n is smaller (larger) than the expected value, i.e., $y = \sum_{i=0}^{n} P(i;\mu)$ for $n < \mu$ and $y = \sum_{i=n}^{\infty} P(i;\mu)$ for $n > \mu$.
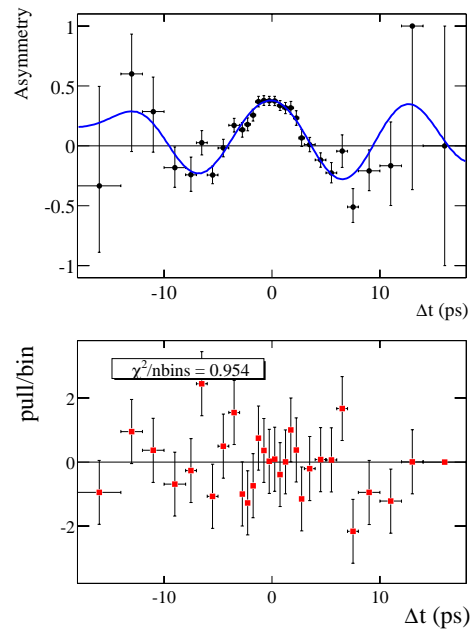
Figure 11.5: The asymmetry plot for signal sample of Data within the peak region $143 < \delta m < 148$ MeV, and the projection of the model fit result from Data. The lower plot shows the bin-by-bin difference normalized by the error estimated from data on each bin.
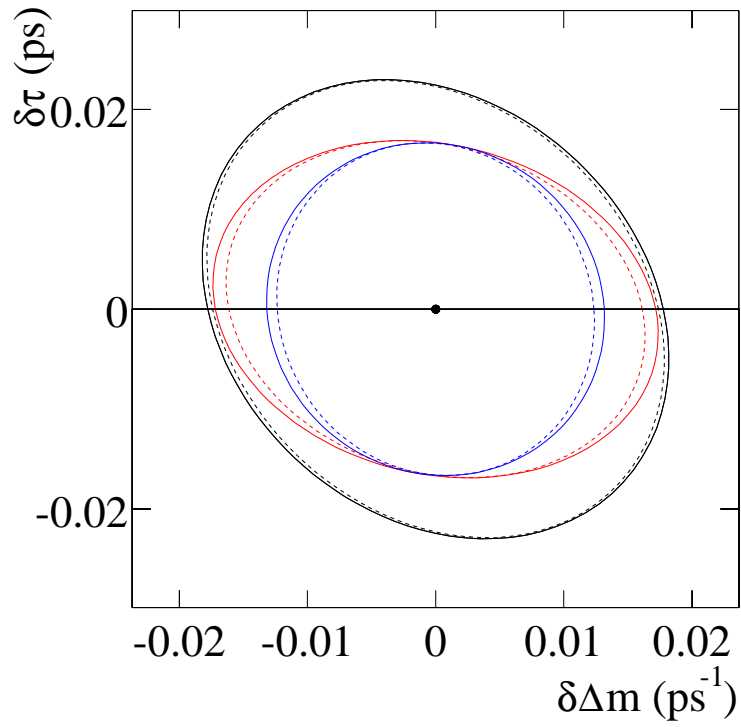
Figure 11.6: Comparison of one-sigma likelihood contours in the Δm_d-τ_{B^0} plane for fits in which different sets of parameters are floating. From the innermost to the outermost ellipse, the floating parameters are (Δm_d, τ_{B^0}), (Δm_d, τ_{B^0}, mistag fractions), (Δm_d, τ_{B^0}, f_{B^+}), (Δm_d, τ_{B^0}, f_{B^+}, mistag fractions), all signal Δt parameters, and the default fit (72 floating parameters).

11.3 Fit Result Summary

The fit results and statistical errors for lifetime and B^0 mixing frequency, and their statistical correlation are:

$$\Delta m_d = 0.512 \pm 0.018 \text{ ps}^{-1} \tag{11.1}$$

$$\tau_{B^0} = 1.545 \ {}^{+0.024}_{-0.023} \text{ ps} \tag{11.2}$$

$$\rho(\Delta m_d, \tau_{B^0}) = -0.22 \tag{11.3}$$

These numbers are before any correction. After corrections, the final numbers are:

$$\Delta m_d = 0.492 \pm 0.018 \text{ ps}^{-1} \tag{11.4}$$

$$\tau_{B^0} = 1.523 \ {}^{+0.024}_{-0.023} \text{ ps} \tag{11.5}$$

$$\rho(\Delta m_d, \tau_{B^0}) = -0.22 \tag{11.6}$$

11.4 Conclusion

Figure 11.7 shows how this measurement compares to other measurements - BaBar, Belle and PDG. This measurement is certainly competitive with other measurements, but the most important aspect of this measurement is the new technology and techniques used to tackle this extremely difficult task.

The $B \rightarrow D^* l \nu$ decay requires more attention to systematic studies which will ultimately help other measurements which will face similar problems in the near future. The new innovations include:
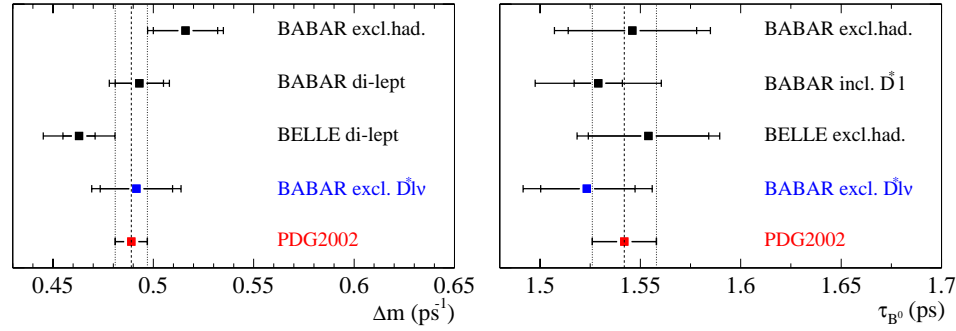
Figure 11.7: Comparison of this measurement (excl. $D^*l\nu$) with others inside BaBar, Belle and the PDG2002.

- Resolution Models: The G+G+G and GExp resolution models were developed with collaboration of several other BaBar analysis groups. Scaling Gaussian bias and widths were introduced along with floating more resolution model parameters. Several BaBar specific problems were resolved including Ktag/per-event-error correlation and the long negative tail in the Δt measurements (long lived charm daughter affecting the vertexing algorithm).

- Including No Tag events: These events don't significantly improve tagging performance, but this measurement has shown that they improve the estimate of resolution parameters which ultimately improves the final statistical error.

- Sub-categories: This measurement has taken the idea of dividing the data into sub-categories to improve the fit to a new level (too far some may say!). As datasets get very large, the systematic errors will dominate over statistical and these ideas will become very important for other groups.

- Improve Fitting Technology: The large datasets and complicated fits quickly bogged down the previous fitting package (RooFitTools). This measurement forced the implementation of several sweeping improvements to the fitting code and created a

new fitting package (RooFitCore/RooFitModels) which made dramatic speed improvements (larger than x10, even more when using lots of sub-categories). The largest improvements came from work addressing the convolution of resolution models and the signal/background PDF's. Other improvements came from more efficient evaluations of Log Likelihood values across sub-categories.

Overall it was an incredibly challenging measurement, but well worth the effort!

Appendix A

Filtered Monte Carlo Sample

The limited amount of generic Monte Carlo is a major concern for this analysis. In order to increase the amount of generic Monte Carlo available, a sample was generated with events which always contain a good lepton with a center of mass momentum greater than 1.0 MeV (not necessarily a good $D^* l\nu$ event, but each event always has at least one good lepton). This filter is 100% efficient for signal events since this analysis requires a good lepton with a center of mass momentum greater than 1.2 MeV. The filter cut rejects about 67% of generic $B^0\overline{B}^0$ and B^+B^- events. This effectively means that each reconstructed event in the filtered sample is equivalent to 3 events in the unfiltered samples.

The filtered and unfiltered samples will have the same yields per unit luminosity of correctly reconstructed events, but the combinatoric, fake and flip lepton control samples could be significantly reduced for the filtered samples. Figure A.1 shows the events found per fb^{-1} between filtered and unfiltered samples. As expected the signal samples are consistent while the fake lepton and combinatoric background samples are greatly reduced. Figure A.2 shows the signal and background shaped for the D^*- mass.

The shapes are completely consistent between the two samples. Finally, the Δt shapes

are shown in figure A.3. The most important point to make is that the filtered sample

doesn't need to match the unfiltered backgrounds to be useful. The combined filtered

and unfiltered sample provides very useful tests for determining the ability of the fitting

algorithm to characterize signal and background events.
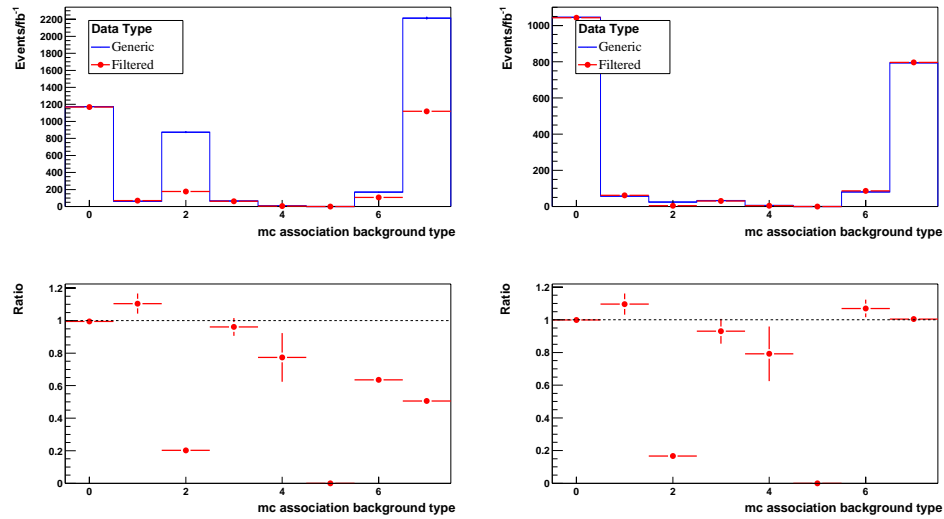
Figure A.1: The number of selected events per fb^{-1} for the entire sample (signal + control samples; left plots) and just the signal sample (right plots). The event type is determined using Monte Carlo truth and is broken down into the following categories: 0=B^0 signal, 1=B^+ event, 2=fake lepton, 3=uncorrelated lepton, 4=cascade lepton, 5=ccbar event, 6=K$\pi\pi^0$ where π^0 missing - just for trouble shooting, 7=combinatoric.
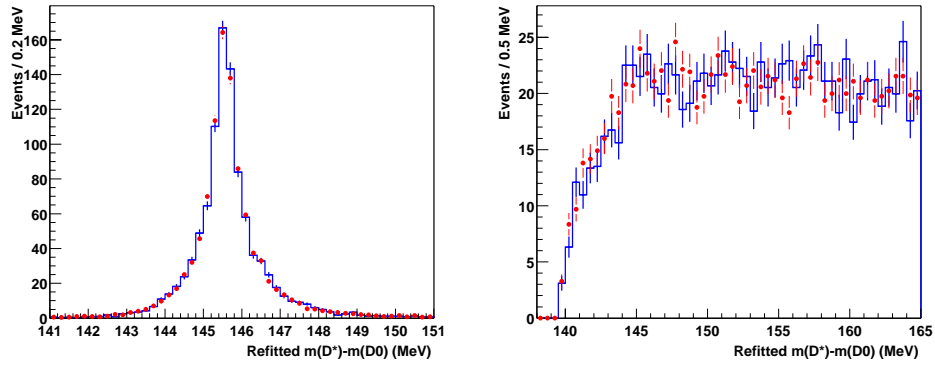
Figure A.2: The D^*- mass shape for correctly reconstructed D^*'s (right) and combinatoric D^*'s (left). All plot normalized to the same area (1000 events).
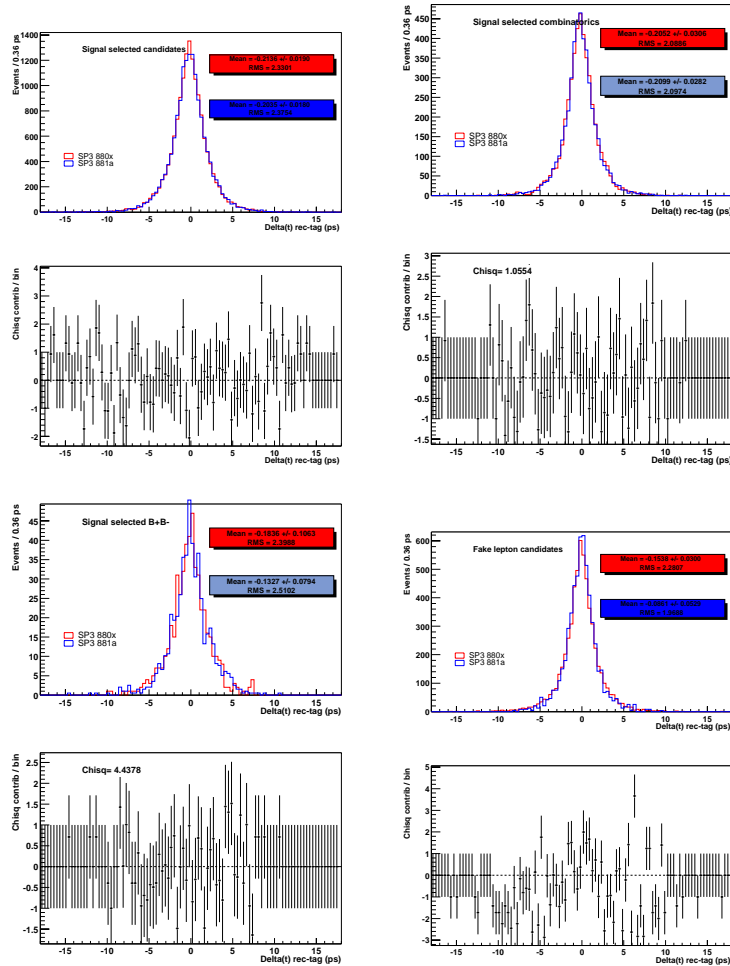
Figure A.3: The Δt plots for signal, combinatoric, charged B and fake samples.

Appendix B

Combinatoric and Peaking Plots

This section contains reference plots related to determining parameters used for the signal probability function, $F(\delta m; \vec{p})$. The first figure show the fit for peak parameters. The next six plot show the fits for background parameters in each of the six peak groups. The final six plots show the makeup of various signal and control samples.

The δm plot parameters can be broken down into two categories, peak parameters and background parameters. There are five peak parameters (the double Gaussian PDF has two widths, two means and a relative fraction) and two background parameters (one for the turn-on portion and one for the long tail). The parameters are described in section 7.2.

As discussed in section 7.2, the peak parameters only depend on SVTDCH(2) $\otimes$ D0mode(3), so these parameters are determined with a simultaneous fit over these six categories shown in figure B.1. For each peak group, the peak parameters are fixed and a new simultaneous fit is performed to determine the two background parameters across the 12 background categories, TagCat(5) $\otimes$ (good/fake lepton) $\oplus$ OnOffRes $\oplus$ AngCut. These plots are shown in figures B.2 through B.7. The remaining six pages show the entire

δm sample broken down by some useful divisions.
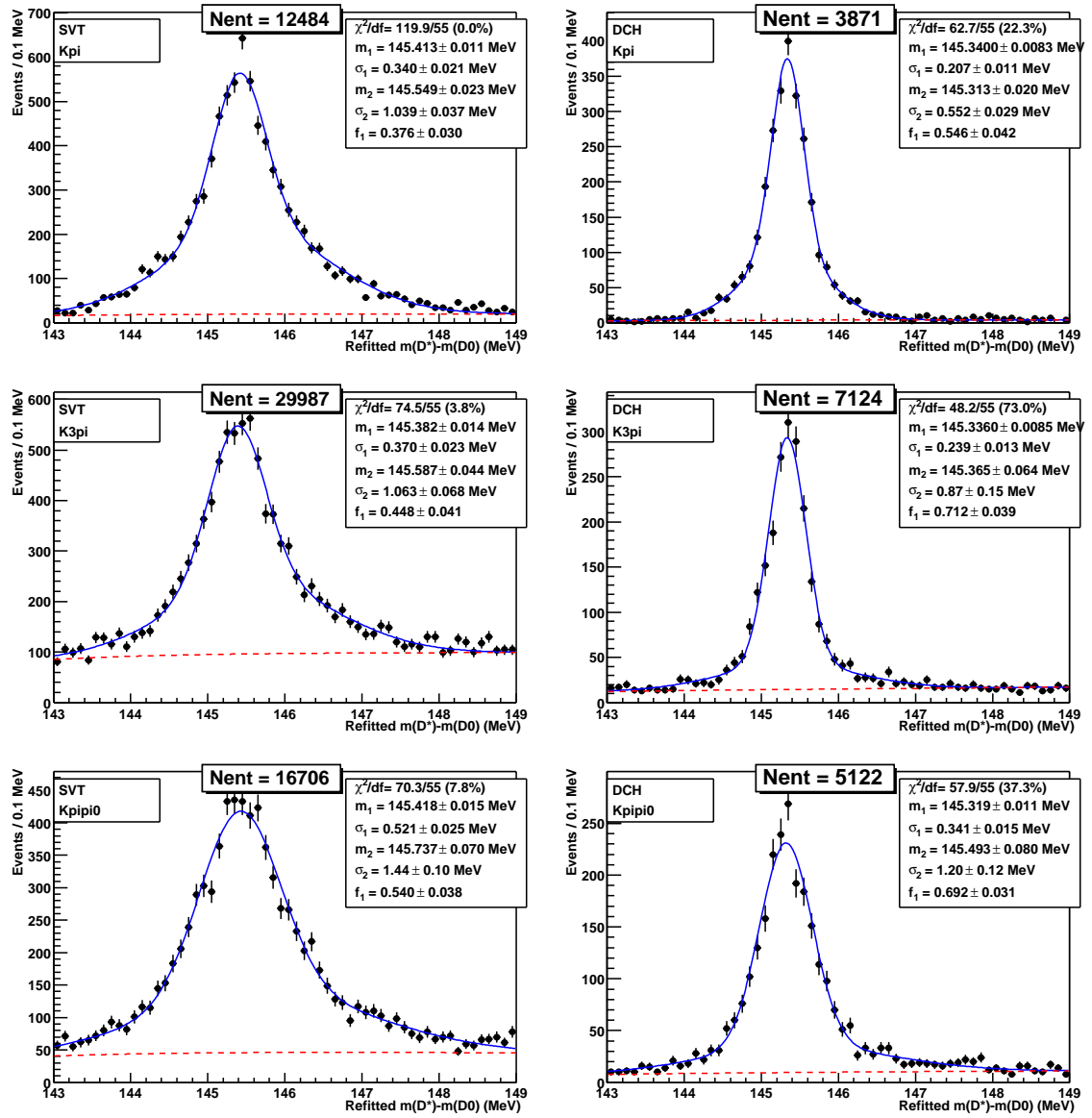
Figure B.1: Results of the simultaneous fit of the 6 peak groups over the peak region 143–149 MeV. The dashed curves show the fitted contributions of combinatoric background.
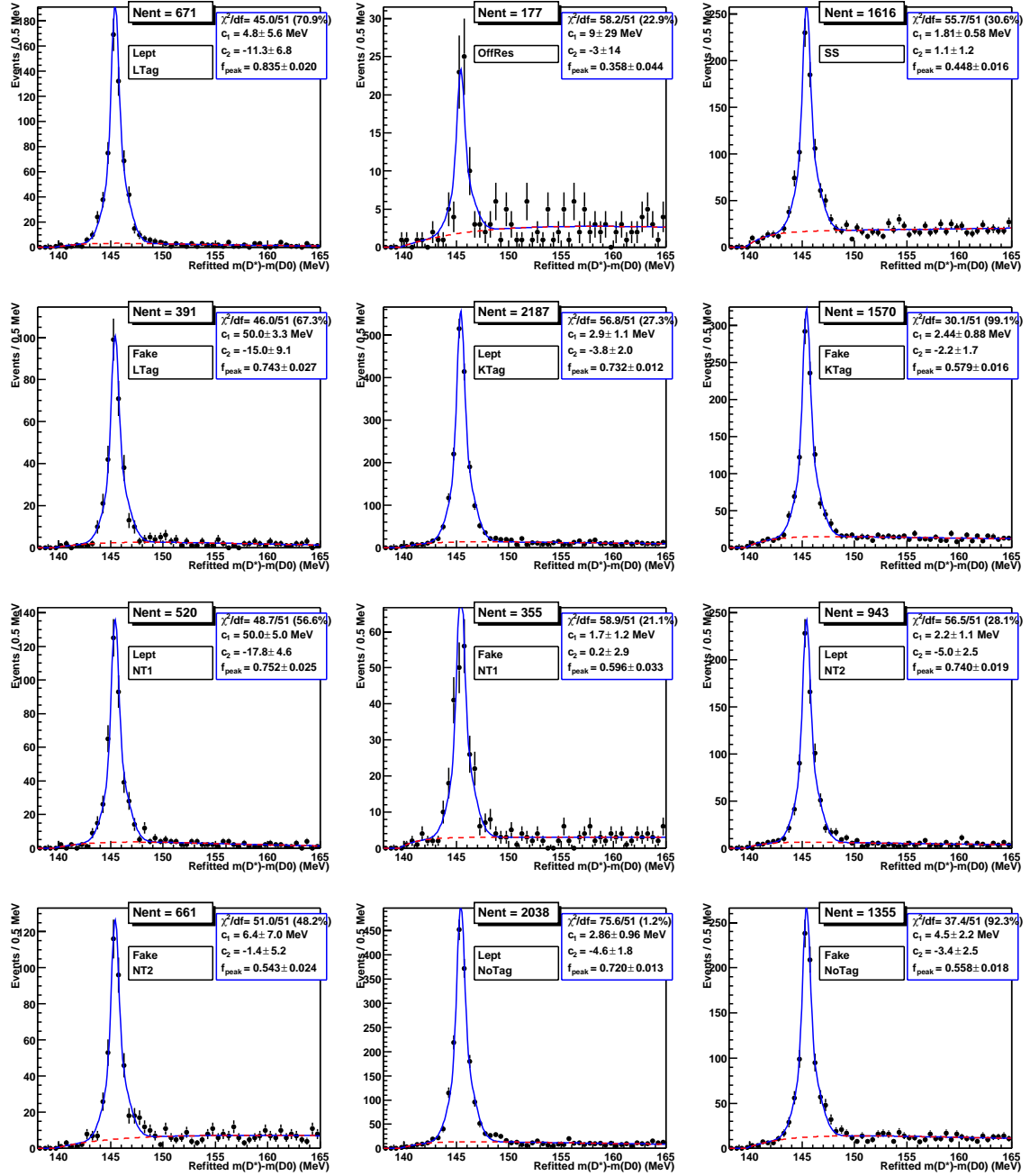
Figure B.2: Results of the simultaneous fit to the peak group $D^0 \to K\pi\times$ SVT only. The dashed curves show the combinatoric background.
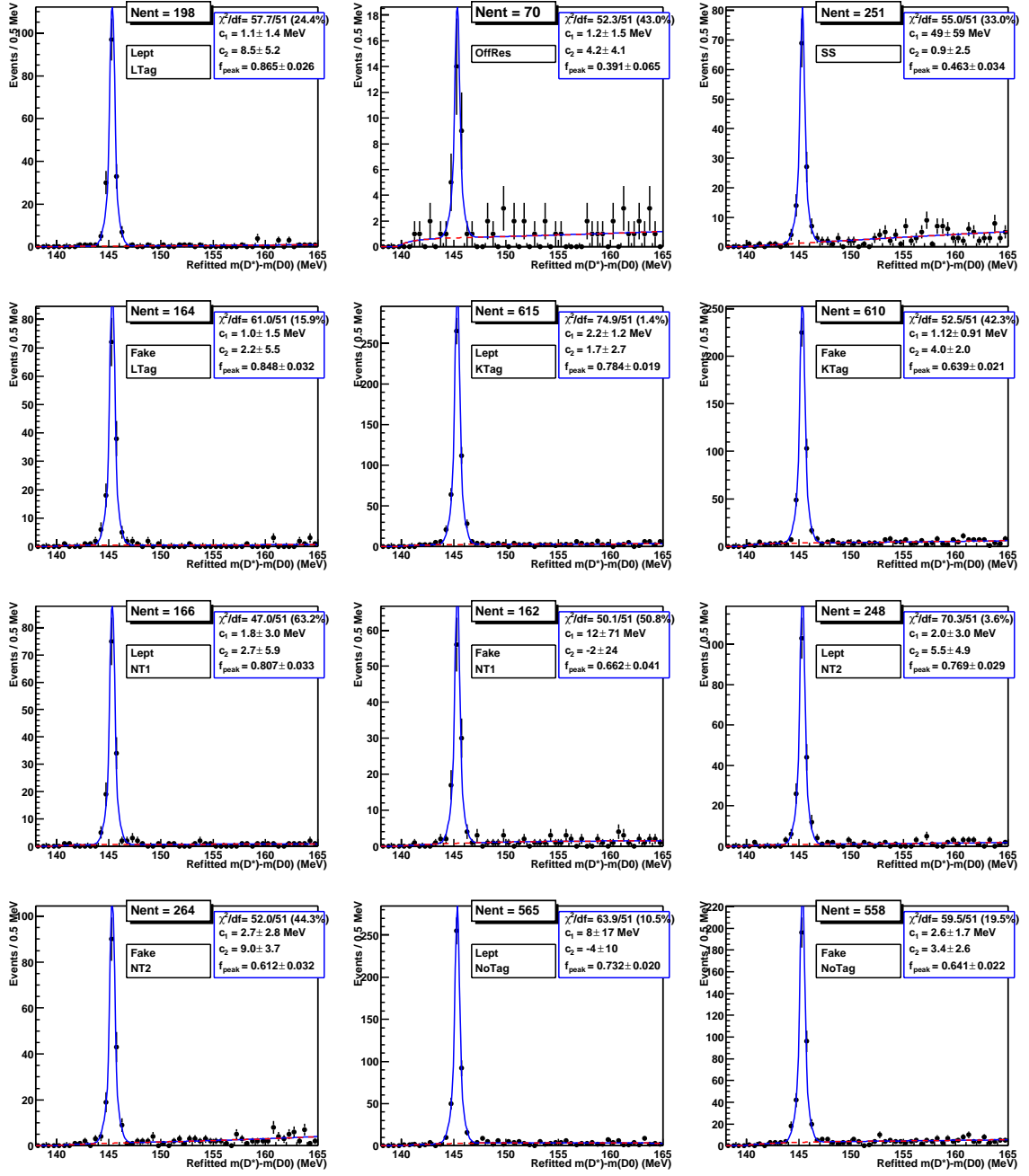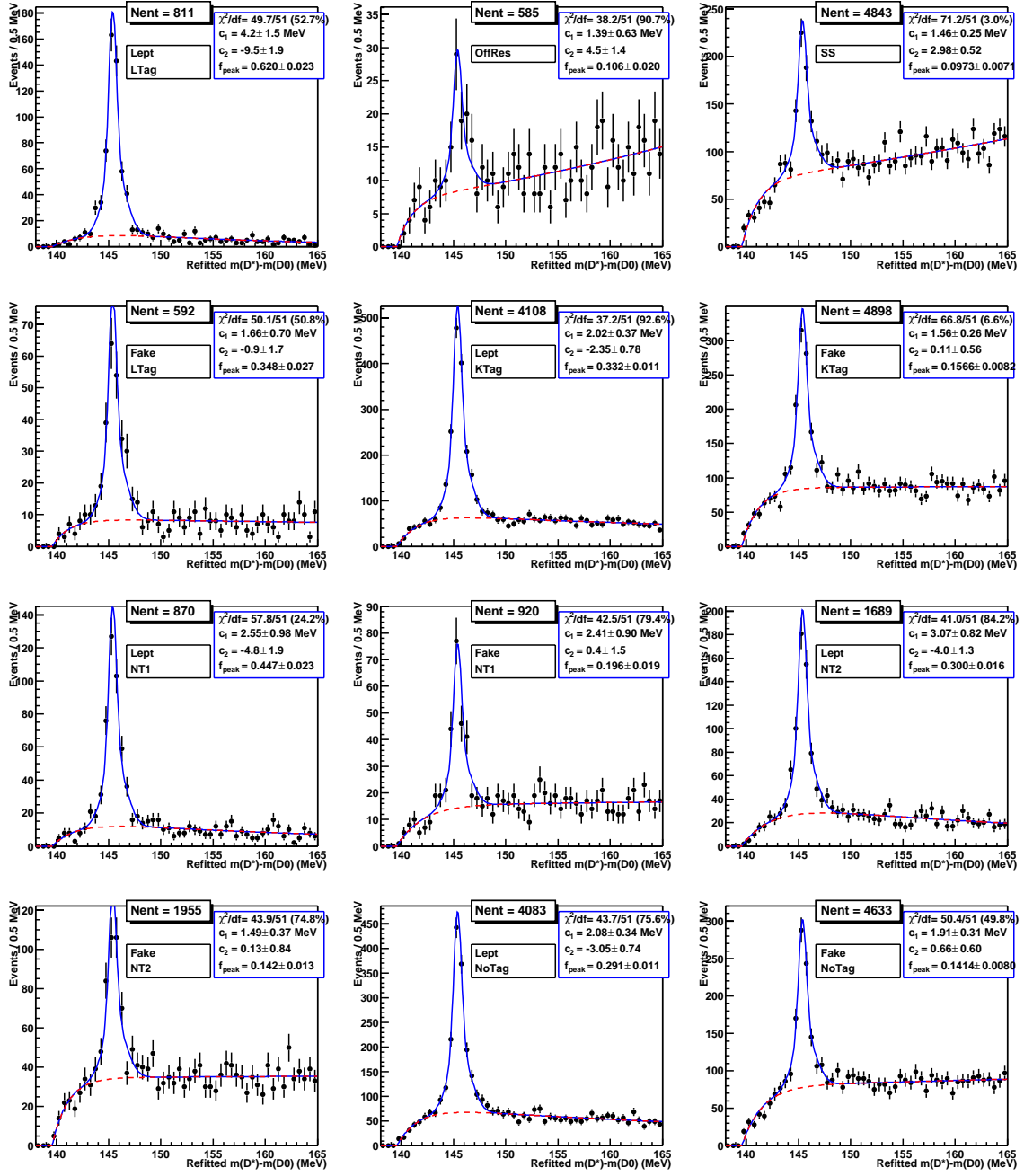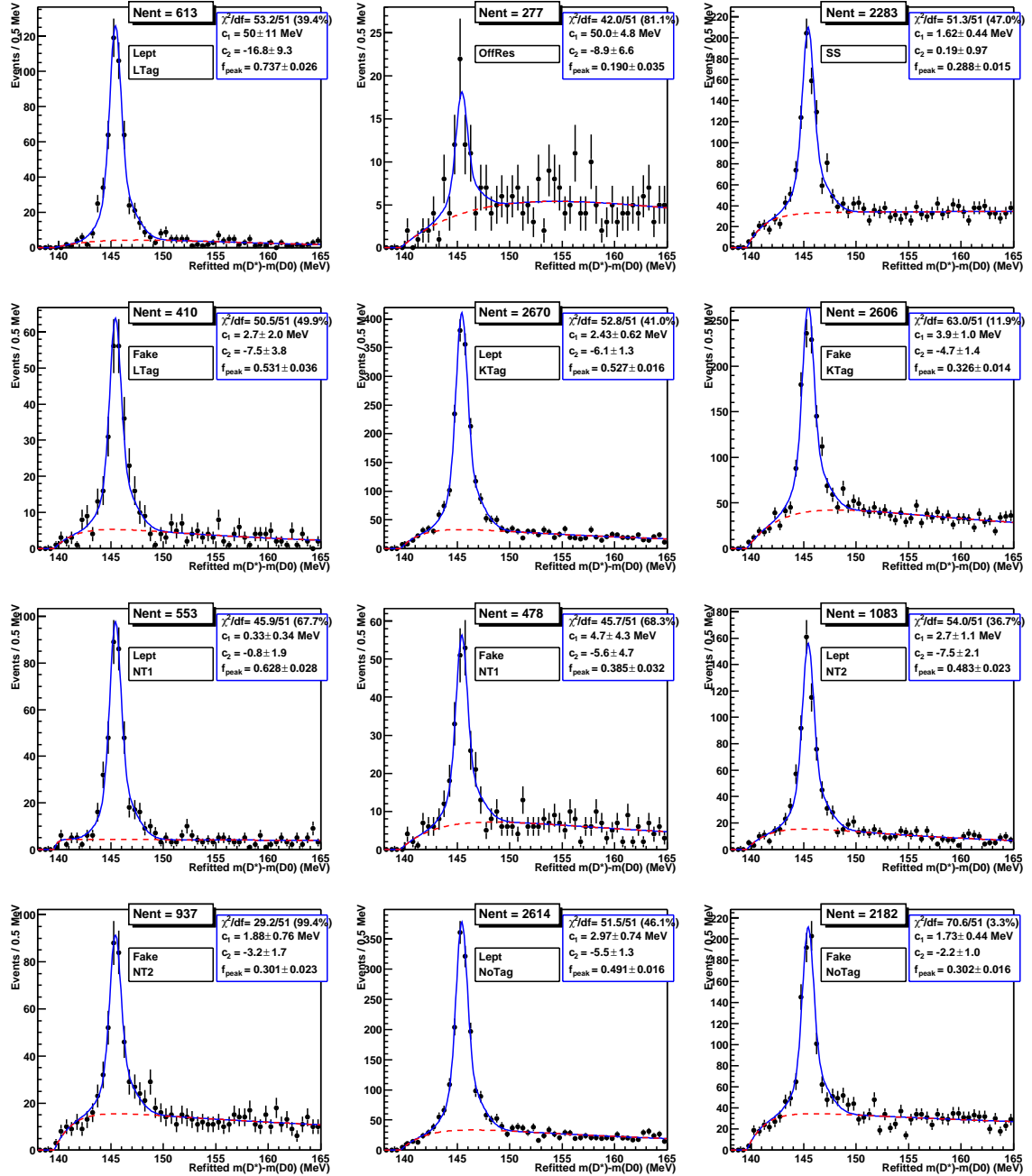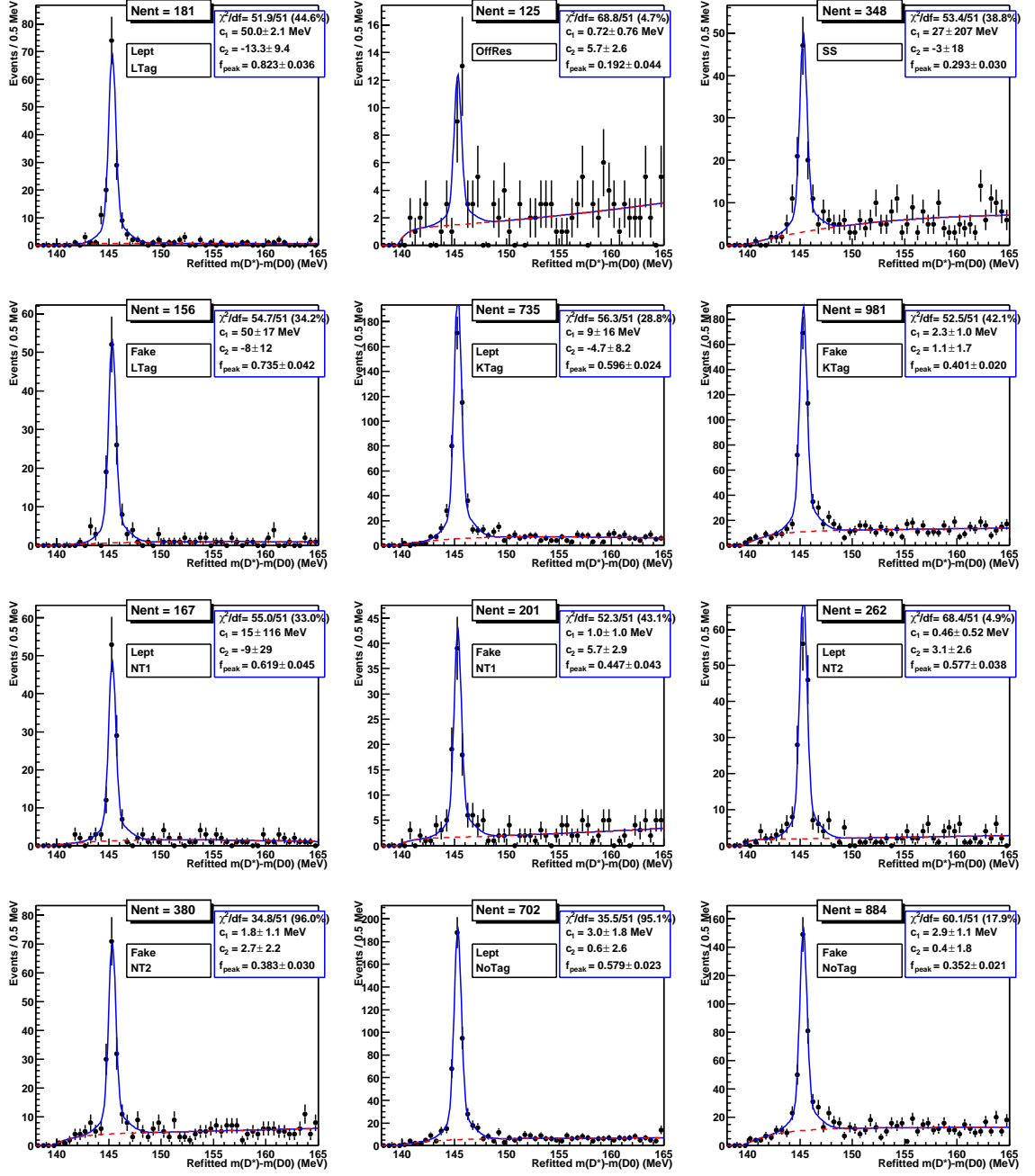
Figure B.3: Results of the simultaneous fit to the peak group $D^0 \to K\pi \times$ SVTDCH. The dashed curves show the combinatoric background.

Figure B.4: Results of the simultaneous fit to the peak group $D^0 \to K\pi\pi\pi$ or $D^0 \to K_S\pi\pi$ $\times$ SVT only. The dashed curves show the combinatoric background.
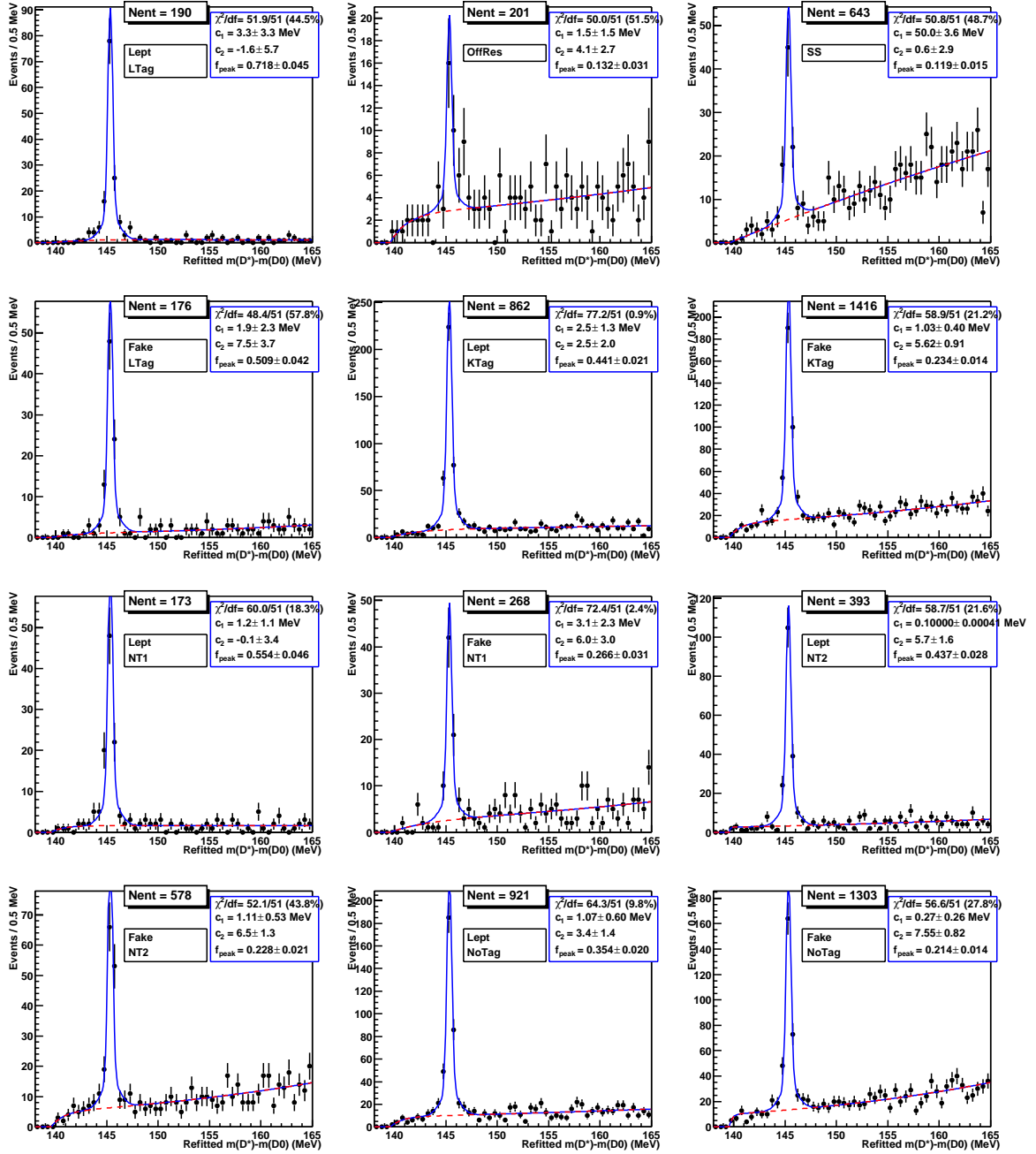
Figure B.5: Results of the simultaneous to the peak group $D^0 \to K\pi\pi\pi$ or $D^0 \to K_S\pi\pi \times$ SVTDCH. The dashed curves show the combinatoric background.

Figure B.6: Results of the simultaneous fit to the peak group $D^0 \to K\pi\pi^0 \times$ SVT only. The dashed curves show the combinatoric background.

Figure B.7: Results of the simultaneous fit to the peak group $D^0 \to K\pi\pi^0 \times$ SVTDCH. The dashed curves show the combinatoric background.
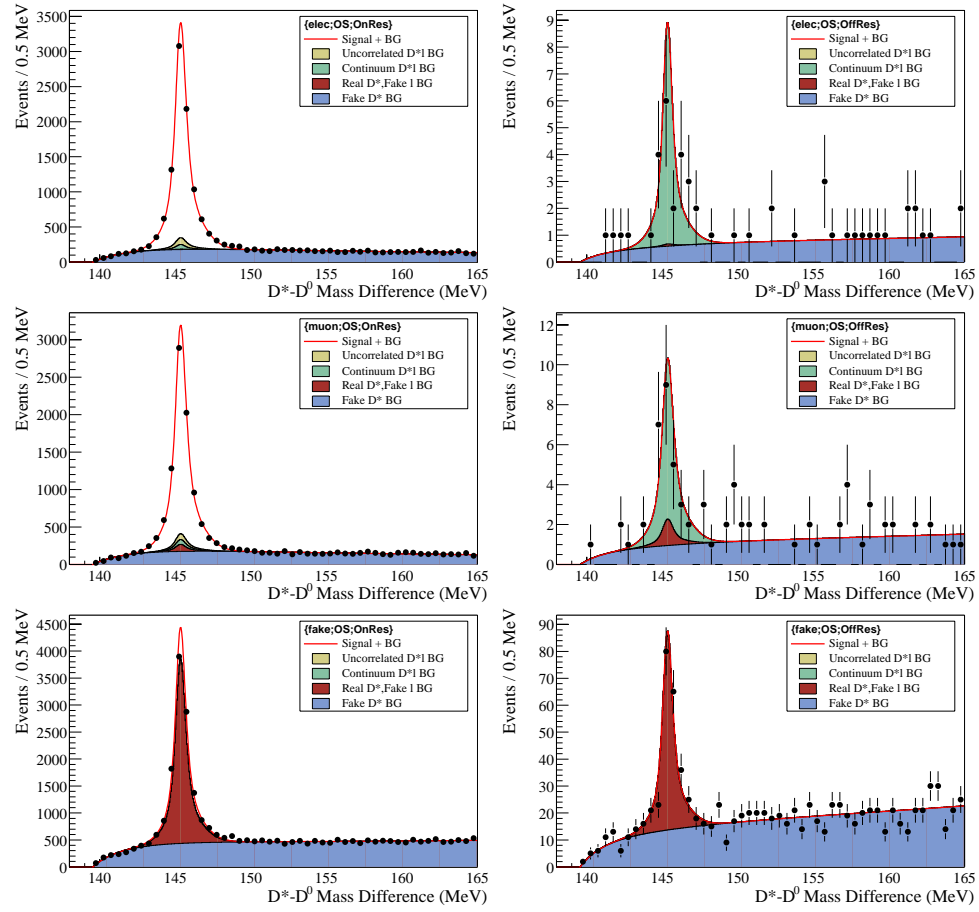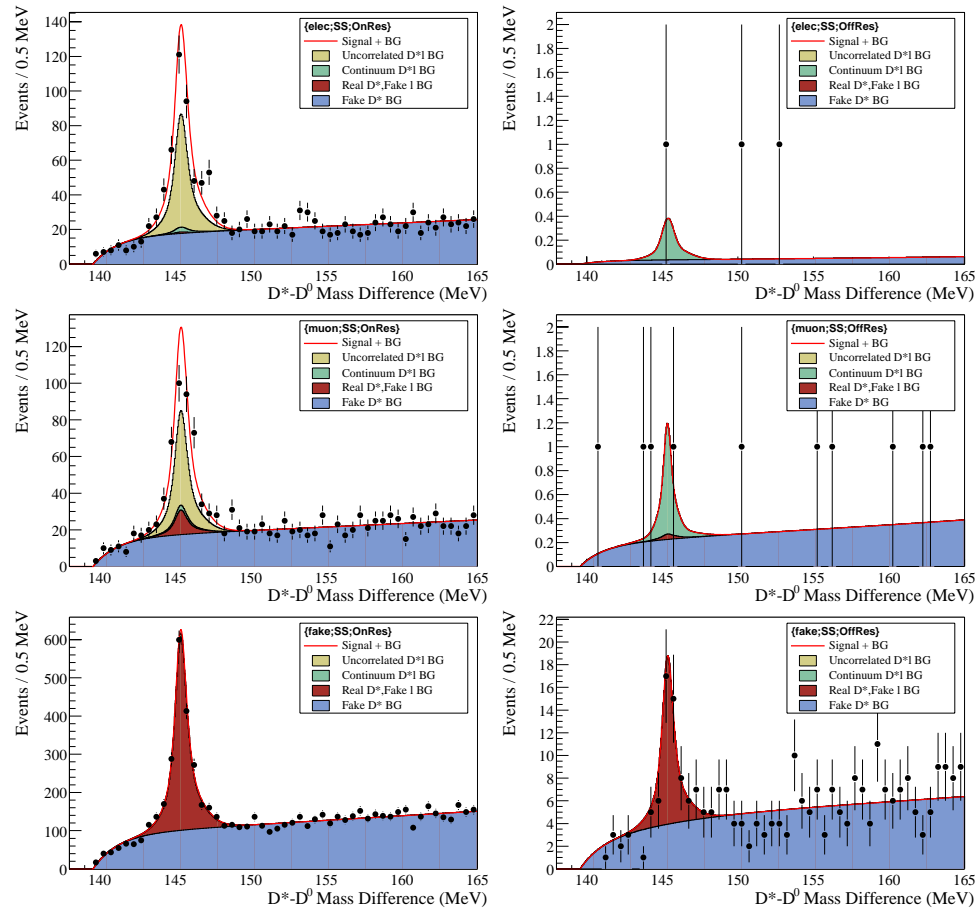
Figure B.8: Uncorrelated $D^*l\nu$, continuum $D^*l\nu$, fake lepton, and fake D^* contributions for Opposite-Side samples. From top to bottom: electron, muon and fake control samples; from left to right: on-resonance and off-resonance.
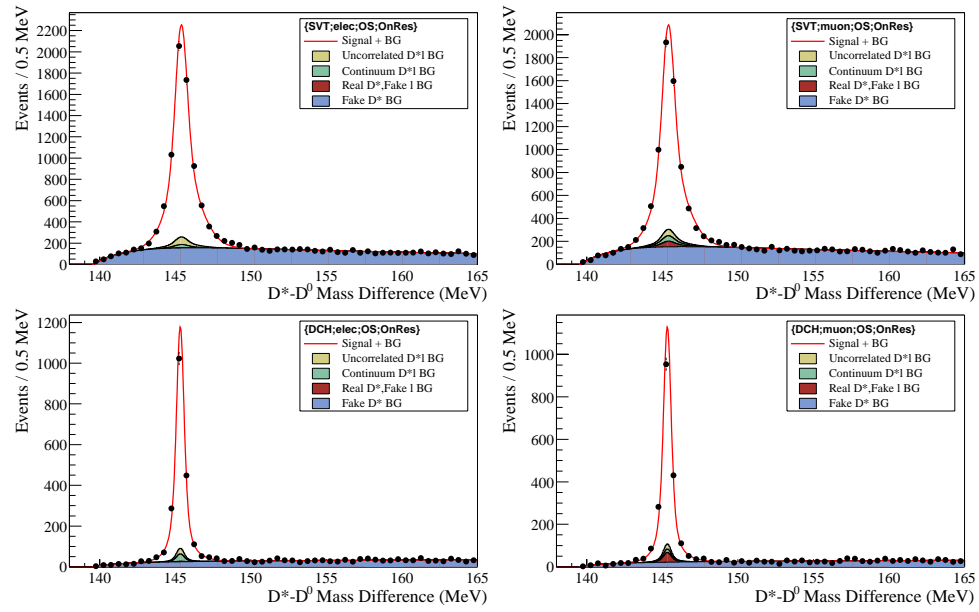
Figure B.9: Uncorrelated $D^*l\nu$, continuum $D^*l\nu$, fake lepton, and fake D^* contributions for Same-Side samples. From top to bottom: electron, muon and fake control samples; from left to right: on-resonance and off-resonance.

Figure B.10: Uncorrelated $D^*l\nu$, continuum $D^*l\nu$, fake lepton, and fake D^* contributions for signal samples. Left: electron; right: muon. Up: SVT-only; bottom: SVT+DCH
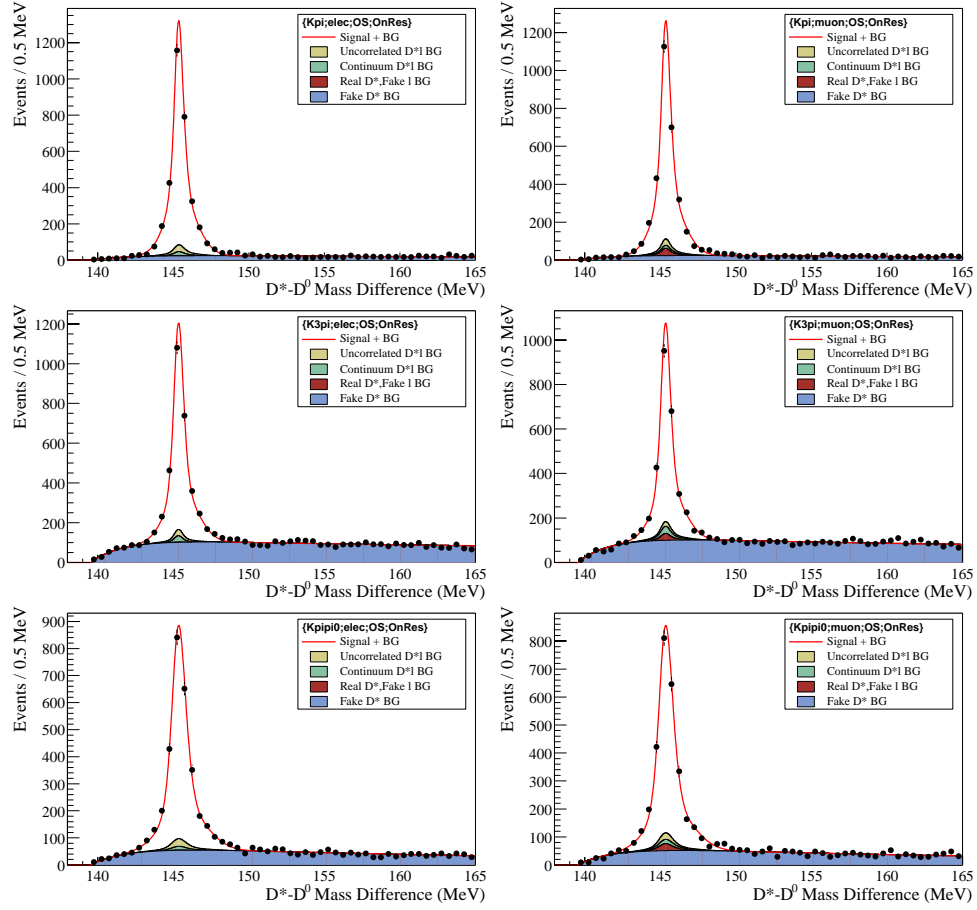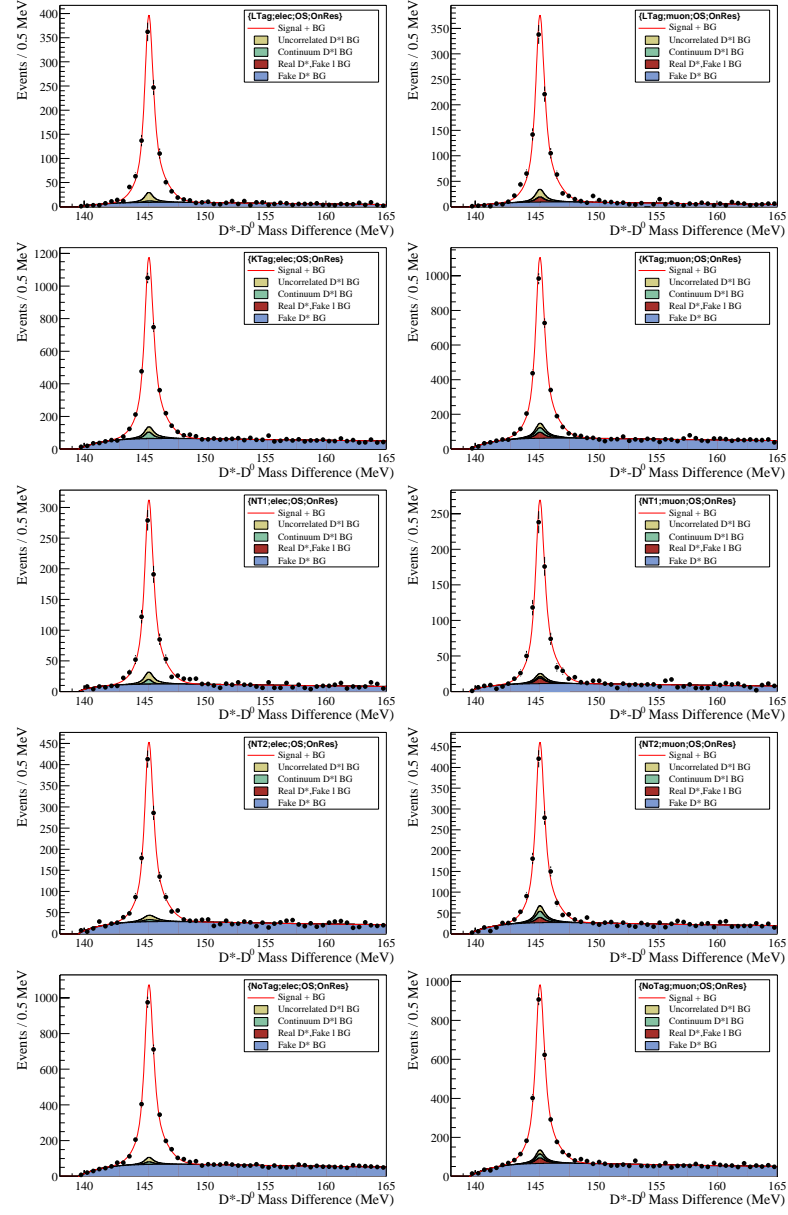
Figure B.11: Uncorrelated $D^* l\nu$, continuum $D^* l\nu$, fake lepton, and fake D^* contributions for signal samples. Left: electron; right: muon. From top to bottom, $D^0 \rightarrow K\pi$, $D^0 \rightarrow K\pi\pi\pi$, and $D^0 \rightarrow K\pi\pi^0$.

Figure B.12: Uncorrelated $D^*l\nu$, continuum $D^*l\nu$, fake lepton, and fake D^* contributions, for signal samples. Left: electron; right: muon. From top to bottom, LeptonTag, KaonTag, NT1Tag, NT2Tag, and NoTag.
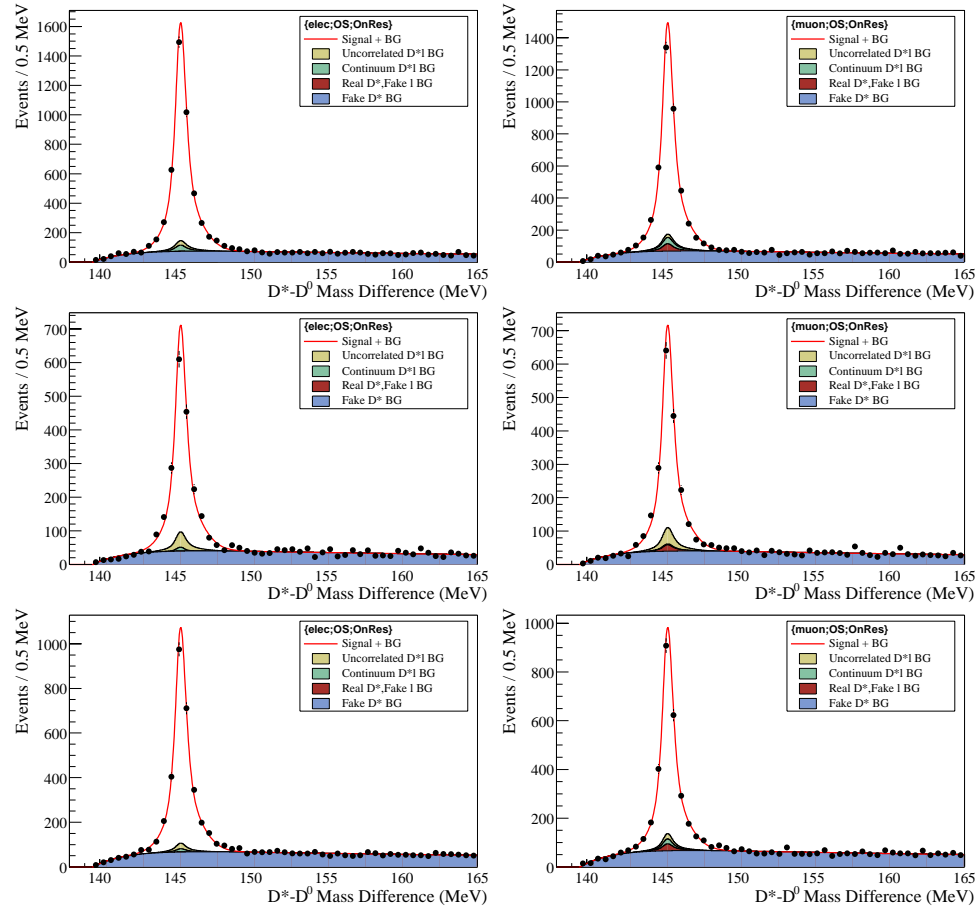
Figure B.13: Peak yields and background contributions for unmixed (top), mixed (middle), and no tagging (bottom) events, for opposite-side on-resonance events only. Electron sample is on the left and muon sample is on the right.

Bibliography

[1] Irish Scientists, "Lord Kelvin", http://members.tripod.com/Ĩrishscientists/listof.htm.

[2] R.N. Cahn, 1996, Rev. Mod. Phys. Vol 68, pg 951.

[3] R. Cahn, Hypernews post, http://babar-hn.slac.stanford.edu:5090/ HyperNews/get/Bmixing/330/2/1.html

[4] Peskin and Schroeder, "An Introduction to Quantum Field Theory", Addison-Westley Publishing, 1997.

[5] F. Halzen and A.D. Martin, "Quarks and Leptons", John Wiley and Sons, 1984.

[6] Review of Particle Physics, D.E. Groom et al., The European Physical Journal C15 (2000)

[7] P. Harrison and H.R. Quinn, "The BaBar Physics Book", SLAC, October 1998.

[8] Y. Nir, "CP Violation In and Beyond the Standard Model", SLAC Summer Institute, 1999.

[9] R.N. Cahn, "B-B Mixing and CP Violation", Preprint of book, private communication, 2001.

[10] R. Cahn, "TagMixZ and its Application to the Analysis of CP Violation", BABAR Analysis Document 17.

[11] D. Kirkby et al, "A User's Guide to the RooFitTools Package for Unbinned Maximum Likelihood Fitting", BABAR Analysis Document 18.

[12] F. Fabozzi et al, "Selection of $D^* l \nu$ Events for B Lifetime and Mixing Analyses", BABAR Analysis Document 34.

[13] N. Roe et al, "Muon Identification in the BaBar Experiment", BABAR Analysis Document 60.

[14] N. Roe et al, "A measurement of B0 B0 oscillation frequency using reconstructed B0 decays", BABAR Analysis Document 81.

[15] Electron PID group, "Cut Based Electron Identification", BABAR Analysis Document 90.

[16] F. Martinez-Vidal et al, "The BaBar Vertexing", BABAR Analysis Document 102.

[17] H. Hu et al, "The NetTagger", BABAR Analysis Document 103.

[18] D. Lange et al, "Supporting document for the Run 1 sin2beta analysis", BABAR Analysis Document 115.

[19] G. Mancinelli et al, "Kaon Selection at the BaBar experiment", BABAR Analysis Document 116.

[20] J. Beringer, S. Plaszczynski, M.H. Schune, A. Roodman, "Cut Based Tagging", BABAR Analysis Document 118.

[21] J. Beringer et al, "B Tagging in BaBar: Status for the $\sin 2\beta$ Journal Publication", BABAR Analysis Document 119.

[22] G. Raven et al, "Measurement of B0 mixing using fully reconstructed hadronic B0 decays", BABAR Analysis Document 125.

[23] D. Kirkby et al, "Measurement of the B^0 Lifetime and Oscillation Frequency Using Fully-Reconstructed Semileptonic Decays", BABAR Analysis Document 137.

[24] J. Stark et al, "Measurements of the charged and neutral B meson lifetimes using fully reconstructed B decays", BABAR Analysis Document 144.

[25] C. Cheng, D. Kirkby, "Semileptonic B Sample for Lifetime and Mixing Analyses", BABAR Analysis Document 147.

[26] C. Borean et al, "A Precise Measurement of B Meson Lifetimes and Mixing with Inclusive Dilepton Events", BABAR Analysis Document 167.

[27] R.N. Cahn, D. Kirkby, "A General Model for Neutral B Decay Time Distributions", BABAR Analysis Document 188.

[28] F. Martinez-Vidal et al, "Vertexing supporting documents for Summer 2001 Conferences", BABAR Analysis Document 254.

[29] http://babar-hn.slac.stanford.edu:5090/HyperNews/get/VertexTools/227.html